

# LECTURE NOTES ON MATHER'S THEORY FOR LAGRANGIAN SYSTEMS

ALFONSO SORRENTINO

## CONTENTS

|                                                                     |    |
|---------------------------------------------------------------------|----|
| 1. Introduction                                                     | 1  |
| 2. Tonelli Lagrangians and Hamiltonians on compact manifolds        | 3  |
| 3. A cartoon example: properties of measures and orbits on KAM tori | 8  |
| 4. Action-minimizing measures: Mather sets                          | 16 |
| 4.A: The symplectic invariance of Mather sets                       | 29 |
| 4.B: An example: the simple pendulum I                              | 33 |
| 4.C: Holonomic measures and generic properties                      | 37 |
| 5. Global action minimizing curves: Aubry and Mañé sets             | 39 |
| 5.A - Some topological and symplectic properties of these sets      | 54 |
| 5.B - An example: the simple pendulum II                            | 55 |
| 5.C - Mather's approach: Peierls' Barrier                           | 58 |
| 6. Weak KAM theory                                                  | 61 |
| 6.A - Regularity of critical subsolutions                           | 68 |
| 6.B - Non-wandering set of the Mañé set                             | 69 |
| References                                                          | 70 |

## 1. INTRODUCTION

The celebrated Kolmogorov-Arnol'd-Moser (or KAM) theorem finally settled the old question concerning the existence of *quasi-periodic* motions for *nearly-integrable* Hamiltonian systems, *i.e.*, Hamiltonian systems that are slight perturbation of an *integrable* one. In the integrable case, in fact, the whole phase space is foliated by invariant *Lagrangian* submanifolds that are diffeomorphic to tori, and on which the dynamics is conjugated to a rigid rotation. These tori are generally called *KAM tori* (see also Definition 3.1). On the other hand, it is natural to ask what happens to such a foliation and to these *stable motions* once the system is perturbed. In 1954 Kolmogorov [30] - and later Arnol'd [1] and Moser [49] in different contexts - proved that, in spite of the generic disappearance of the invariant tori filled by periodic orbits, already pointed out by Henri Poincaré, for small perturbations of an integrable system it is still possible to find invariant Lagrangian tori corresponding to “*strongly non-resonant*”, *i.e.*, *Diophantine*, rotation vectors. This result, commonly referred to as *KAM theorem*, from the initials of the three main pioneers, besides opening the way to a new understanding of the nature of Hamiltonian systems and their stable motions, contributed to raise new interesting questions, for instance about the destiny of the *stable motions* (orbits on KAM tori) that are destroyed by effect of the perturbation or about the possibility of extending these results to

---

*Date:* November 3, 2010.

PRELIMINARY VERSION: PLEASE SEND ALL COMMENTS AND TYPOS TO THE AUTHOR.

systems that are not close to any integrable one.

An answer to the first question did not take long to arrive. In 1964 V. I. Arnol'd [2] constructed an example of a perturbed integrable system, in which *unstable orbits* - resulting from the breaking of unperturbed KAM tori - coexist with the stable motions drawn by KAM theorem. This striking, and somehow unexpected, phenomenon, yet quite far from being completely understood, is nowadays called *Arnol'd diffusion*. This new insight led to a change of perspective and in order to make sense of the complex balance between stable and unstable motions that was looming out, new approaches needed to be exploited. Amongst these, variational methods turned out to be particularly successful. Mostly inspired by the so-called *least action principle*, a sort of widely accepted “thriftiness” of Nature in all its actions, they seemed to provide the natural setting to get over the local view given by the analytical methods and make towards a global understanding of the dynamics.

*Aubry-Mather theory* represents probably one of the biggest triumphs in this direction. Developed independently by Serge Aubry [3] and John Mather [38] in the eighties, this novel approach to the study of the dynamics of *twist diffeomorphisms of the annulus* (which correspond to Poincaré maps of 1-dimensional Hamiltonian systems [49]) pointed out the existence of many *action-minimizing sets*, which in some sense generalize invariant rotational curves and that always exist, even after rotational curves are destroyed. Besides providing a detailed structure theory for these new sets, this powerful approach yielded to a better understanding of the destiny of invariant rotational curves and to the construction of interesting chaotic orbits as a result of their destruction [40, 27, 43].

Motivated by these achievements, John Mather [42, 44] - and later Ricardo Mañé [35, 14] and Albert Fathi [22] in different ways - developed a generalization of this theory to higher dimensional systems. Positive definite superlinear Lagrangians on compact manifolds, also called *Tonelli Lagrangians* (see Definition 2.1), were the appropriate setting to work in. Under these conditions, in fact, it is possible to prove the existence of interesting invariant (action-minimizing) sets, known as *Mather*, *Aubry* and *Mañé* sets, which generalize KAM tori, and which continue to exist even after KAM tori's disappearance or when it does not make sense to speak of them (for example when the system is “far” from any integral one).

Let us remark that these tools revealed also quite promising in the construction of *chaotic orbits*, such as for instance *connecting orbits* among the above-mentioned invariant sets [44, 5, 17]. Therefore they set high hopes on the possibility of proving the generic existence of Arnold diffusion in nearly integrable Hamiltonian systems [45]. However, differently from the case of twist diffeomorphisms, the situation turns out to be more complicated, due to a general lack of information on the topological structure of these action-minimizing sets. These sets, in fact, play a twofold role. Whereas on the one hand they may provide an obstruction to the existence of “diffusing orbits”, on the other hand their topological structure plays a fundamental role in the variational methods that have been developed for the construction of orbits with “prescribed” behaviors. We shall not enter further into the discussion of this problematic, but we refer the interested readers to [5, 9, 17, 45, 48].

In these lecture notes we shall try to provide a brief, but hopefully comprehensive introduction to Mather's theory for Lagrangian systems and its subsequent developments by Ricardo Mañé and Albert Fathi. We shall consider only the autonomous case (*i.e.*, no dependence on time in the Lagrangian and Hamiltonian). This choice has been made only to make the discussion easier and to avoid some technical issues that would be otherwise involved. However, all the theory that we are going to describe can be generalized, with some “small” modifications, to the

non-autonomous time-periodic case. Along our discussion, in order to draw the most complete picture of the theory, we shall point out and discuss such differences and the needed modifications.

The sections will be organized as follows:

- Section 2: we shall introduce Tonelli Lagrangians and Hamiltonians on compact manifolds and discuss their properties and some examples.
- Section 3: before entering into the description of Mather's theory, we shall discuss a cartoon example, namely, the properties of invariant probability measures and orbits on KAM tori. This will prepare the ground for understanding the ideas behind Mather's work, as well Mañé and Fathi's ones.
- Section 4: we shall discuss the notion of action minimizing measures and introduce the first family of invariant sets: the *Mather sets*.
- Section 5: we shall discuss the notion of action minimizing orbits and introduce other two families of invariant sets: the *Aubry* and *Mane sets*.
- Section 6: we shall discuss Fathi's *Weak KAM theory* and its relation to Mather and Mañé's works.
- Some Addenda to the single sections with some complimentary material will be provided along the way.

*Acknowledgements.* These lectures were delivered at Università degli Studi di Napoli "Federico II" (April 2009, thematic program "*New connections between dynamical systems and Hamiltonian PDEs*") and at Universitat Politècnica de Catalunya (June 2010, summer school "*Jornades d'introducció als sistemes dinàmics i a les EDP's*"). I am very grateful to, respectively, Massimiliano Berti, Michela Procesi, Vittorio Coti-Zelati and Xavier Cabré, Amadeu Delshams, Maria del Mar Gonzales, Tere M. Seara, for their kind invitation. I would also like to thank all participants to the courses for their useful comments, stimulating suggestions and careful feedbacks, in particular Will Merry, Joana Dos Santos and Rodrigo Treviño for a careful reading of a first draft of these notes.

A special acknowledgement must go to John Mather, Albert Fathi and Patrick Bernard, from whom I learnt most of the material here collected and much beyond. I would also like to thank all other people that have contributed, in different ways and at different times, to the realization of this project: Luigi Chierchia, Gonzalo Contreras, Rafael de la Llave, Daniel Massart, Gabriel Paternain and many others.

## 2. TONELLI LAGRANGIANS AND HAMILTONIANS ON COMPACT MANIFOLDS

In this section we want to introduce the basic setting that we shall be considering hereafter. Let  $M$  be a compact and connected smooth manifold without boundary. Denote by  $TM$  its tangent bundle and  $T^*M$  the cotangent one. A point of  $TM$  will be denoted by  $(x, v)$ , where  $x \in M$  and  $v \in T_x M$ , and a point of  $T^*M$  by  $(x, p)$ , where  $p \in T_x^* M$  is a linear form on the vector space  $T_x M$ . Let us fix a Riemannian metric  $g$  on it and denote by  $d$  the induced metric on  $M$ ; let  $\|\cdot\|_x$  be the norm induced by  $g$  on  $T_x M$ ; we shall use the same notation for the norm induced on  $T_x^* M$ .

We shall consider functions  $L : TM \rightarrow \mathbb{R}$  of class  $C^2$ , which are called *Lagrangians*. Associated to each Lagrangian, there is a flow on  $TM$  called the *Euler-Lagrange flow*, defined as follows. Let us consider the action functional  $A_L$  from

the space of continuous piecewise  $C^1$  curves  $\gamma : [a, b] \rightarrow M$ , with  $a \leq b$ , defined by:

$$A_L(\gamma) := \int_a^b L(\gamma(t), \dot{\gamma}(t)) dt.$$

Curves that extremize this functional among all curves with the same end-points are solutions of the *Euler-Lagrange equation*:

$$(1) \quad \frac{d}{dt} \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)) = \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t)) \quad \forall t \in [a, b].$$

Observe that this equation is equivalent to

$$\frac{\partial^2 L}{\partial v^2}(\gamma(t), \dot{\gamma}(t)) \ddot{\gamma}(t) = \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t)) - \frac{\partial^2 L}{\partial v \partial x}(\gamma(t), \dot{\gamma}(t)) \dot{\gamma}(t),$$

therefore, if the second partial vertical derivative  $\partial^2 L / \partial v^2(x, v)$  is non-degenerate at all points of  $TM$ , we can solve for  $\ddot{\gamma}(t)$ . This condition

$$\det \frac{\partial^2 L}{\partial v^2} \neq 0$$

is called *Legendre condition* and allows one to define a vector field  $X_L$  on  $TM$ , such that the solutions of  $\ddot{\gamma}(t) = X_L(\gamma(t), \dot{\gamma}(t))$  are precisely the curves satisfying the Euler-Lagrange equation. This vector field  $X_L$  is called the *Euler-Lagrange vector field* and its flow  $\Phi_t^L$  is the *Euler-Lagrange flow* associated to  $L$ . It turns out that  $\Phi_t^L$  is  $C^1$  even if  $L$  is only  $C^2$  (see Remark 2.4).

**Definition 2.1 (Tonelli Lagrangian).** *A function  $L : TM \rightarrow \mathbb{R}$  is called a Tonelli Lagrangian if:*

- i)  $L \in C^2(TM)$ ;
- ii)  $L$  is strictly convex in the fibers, in the  $C^2$  sense, i.e., the second partial vertical derivative  $\partial^2 L / \partial v^2(x, v)$  is positive definite, as a quadratic form, for all  $(x, v)$ ;
- iii)  $L$  is superlinear in each fiber, i.e.,

$$\lim_{\|v\|_x \rightarrow +\infty} \frac{L(x, v)}{\|v\|_x} = +\infty.$$

*This condition is equivalent to ask that for each  $A \in \mathbb{R}$  there exists  $B(A) \in \mathbb{R}$  such that*

$$L(x, v) \geq A\|v\| - B(A) \quad \forall (x, v) \in TM.$$

Observe that since the manifold is compact, then condition *iii*) is independent of the choice of the Riemannian metric  $g$ .

**Remark 2.2.** More generally, one can consider the case of a *time-periodic Tonelli Lagrangian*  $L : TM \times \mathbb{T} \rightarrow \mathbb{R}$  (also called *non-autonomous case*), as it was originally done by John Mather [42]. In fact, as it was pointed out by Jürgen Moser, this was the right setting to generalize Aubry and Mather's results for twist maps to higher dimensions; in fact, every twist map can be seen as the time one map associated to the flow of a periodic Tonelli Lagrangian on the one dimensional torus (see for instance [49]). In this case, a further condition on the Lagrangian is needed:

- iv) *The Euler-Lagrange flow is complete, i.e., every maximal integral curve of the vector field  $X_L$  has all  $\mathbb{R}$  as its domain of definition.*

In the non-autonomous case, in fact, this condition is necessary in order to have that action-minimizing curves (or *Tonelli minimizers*, see section 5) satisfy the Euler-Lagrange equation. Without such an assumption Ball and Mizel [4] have constructed an example of Tonelli minimizers that are not  $C^1$  and therefore are not

solutions of the Euler-Lagrange flow. The role of the completeness hypothesis can be explained as follows. It is possible to prove, under the above conditions, that action minimizing curves not only exist and are absolutely continuous, but they are  $C^1$  on an open and dense full measure subset of the interval in which they are defined. It is possible to check that they satisfy the Euler-Lagrange equation on this set, while their velocity goes to infinity on the exceptional set on which they are not  $C^1$ . Asking the flow to be complete, therefore, implies that Tonelli minimizers are  $C^1$  everywhere and that they are actual solutions of the Euler-Lagrange equation. A sufficient condition for the completeness of the Euler-Lagrange flow, for example, can be expressed in terms of a growth condition for  $\partial L/\partial t$ :

$$-\frac{\partial L}{\partial t}(x, v, t) \leq C \left( 1 + \frac{\partial L}{\partial v}(x, v, t) \cdot v - L(x, v, t) \right) \quad \forall (x, v, t) \in TM \times \mathbb{T}.$$

### Examples of Tonelli Lagrangians.

- **Riemannian Lagrangians.** Given a Riemannian metric  $g$  on  $TM$ , the *Riemannian Lagrangian* on  $(M, g)$  is given by the *Kinetic energy*:

$$L(x, v) = \frac{1}{2} \|v\|_x^2.$$

Its Euler-Lagrange equation is the equation of the geodesics of  $g$ :

$$\frac{D}{dt} \dot{x} \equiv 0,$$

and its Euler-Lagrange flow coincides with the geodesic flow.

- **Mechanical Lagrangians.** These Lagrangians play a key-role in the study of classical mechanics. They are given by the sum of the kinetic energy and a *potential*  $U : M \longrightarrow \mathbb{R}$ :

$$L(x, v) = \frac{1}{2} \|v\|_x^2 + U(x).$$

The associated Euler-Lagrange equation is given by:

$$\frac{D}{dt} \dot{x} = \nabla U(x),$$

where  $\nabla U$  is the gradient of  $U$  with respect to the Riemannian metric  $g$ , *i.e.*,

$$d_x U \cdot v = \langle \nabla U(x), v \rangle_x \quad \forall (x, v) \in TM.$$

- **Mañé's Lagrangians.** This is a particular class of Tonelli Lagrangians, introduced by Ricardo Mañé in [33] (see also [23]). If  $X$  is a  $C^k$  vector field on  $M$ , with  $k \geq 2$ , one can embed its flow  $\varphi_t^X$  into the Euler-Lagrange flow associated to a certain Lagrangian, namely

$$L_X(x, v) = \frac{1}{2} \|v - X(x)\|_x^2.$$

It is quite easy to check that the integral curves of the vector field  $X$  are solutions to the Euler-Lagrange equation. In particular, the Euler-Lagrange flow  $\Phi_t^{L_X}$  restricted to  $\text{Graph}(X) = \{(x, X(x)), x \in M\}$  (that is clearly invariant) is conjugated to the flow of  $X$  on  $M$  and the conjugation is given by  $\pi|_{\text{Graph}(X)}$ , where  $\pi : TM \rightarrow M$  is the canonical projection. In other words, the following diagram commutes:

$$\begin{array}{ccc} \text{Graph}(X) & \xrightarrow{\Phi_t^{L_X}} & \text{Graph}(X) \\ \pi \downarrow & & \downarrow \pi \\ M & \xrightarrow{\varphi_t^X} & M \end{array}$$

that is, for every  $x \in M$  and every  $t \in \mathbb{R}$ ,  $\Phi_t^{L_X}(x, X(x)) = (\gamma_x^X(t), \dot{\gamma}_x^X(t))$ , where  $\gamma_x^X(t) = \varphi_t^X(x)$ .

In the study of classical dynamics, it turns often very useful to consider the associated *Hamiltonian system*, which is defined on the cotangent space  $T^*M$ . Let us describe how to define this new system and what is its relation with the Lagrangian one.

A standard tool in the study of convex functions is the so-called *Fenchel transform*, which allows one to transform functions on a vector space into functions on the dual space (see for instance [22, 51] for excellent introductions to the topic). Given a Lagrangian  $L$ , we can define the associated *Hamiltonian*, as its Fenchel transform (or *Fenchel-Legendre transform*):

$$\begin{aligned} H : T^*M &\longrightarrow \mathbb{R} \\ (x, p) &\longmapsto \sup_{v \in T_x M} \{ \langle p, v \rangle_x - L(x, v) \} \end{aligned}$$

where  $\langle \cdot, \cdot \rangle_x$  denotes the canonical pairing between the tangent and cotangent space.

If  $L$  is a Tonelli Lagrangian, one can easily prove that  $H$  is finite everywhere (as a consequence of the superlinearity of  $L$ ),  $C^2$ , superlinear and strictly convex in each fiber (in the  $C^2$  sense). Such a Hamiltonian is called *Tonelli* (or *optical*) *Hamiltonian*.

**Definition 2.3 (Tonelli Hamiltonian).** *A function  $H : T^*M \longrightarrow \mathbb{R}$  is called a Tonelli (or optical) Hamiltonian if:*

- i)  $H$  is of class  $C^2$ ;
- ii)  $H$  is strictly convex in each fiber in the  $C^2$  sense, i.e., the second partial vertical derivative  $\partial^2 H / \partial p^2(x, p)$  is positive definite, as a quadratic form, for any  $(x, p) \in T^*M$ ;
- iii)  $H$  is superlinear in each fiber, i.e.,

$$\lim_{\|p\|_x \rightarrow +\infty} \frac{H(x, p)}{\|p\|_x} = +\infty.$$

### Examples of Tonelli Hamiltonians.

Let us see what are the Hamiltonians associated to the Tonelli Lagrangians that we have introduced in the previous examples.

- **Riemannian Hamiltonians.** If  $L(x, v) = \frac{1}{2}\|v\|_x^2$  is the Riemannian Lagrangian associated to a Riemannian metric  $g$  on  $M$ , the corresponding Hamiltonian will be

$$H(x, p) = \frac{1}{2}\|p\|_x^2,$$

where  $\|\cdot\|$  represents - in this last expression - the induced norm on the cotangent space  $T^*M$ .

- **Mechanical Hamiltonians.** If  $L(x, v) = \frac{1}{2}\|v\|_x^2 + U(x)$  is a mechanical Lagrangian, the associated Hamiltonian is:

$$H(x, p) = \frac{1}{2}\|p\|_x^2 - U(x),$$

that it is sometime referred to as *mechanical energy*.

- **Mañé's Hamiltonians.** If  $X$  is a  $C^k$  vector field on  $M$ , with  $k \geq 2$ , and  $L_X(x, v) = \|v - X(x)\|_x^2$  is the associated Mañé Lagrangian, one can check that the corresponding Hamiltonian is given by:

$$H(x, p) = \frac{1}{2} \|p\|_x^2 + \langle p, X(x) \rangle.$$

Given a Hamiltonian one can consider the associated *Hamiltonian flow*  $\Phi_t^H$  on  $T^*M$ . In local coordinates, this flow can be expressed in terms of the so-called *Hamilton's equations*:

$$(2) \quad \begin{cases} \dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t)) \\ \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t)). \end{cases}$$

We shall denote by  $X_H(x, p) := \left( \frac{\partial H}{\partial p}(x, p), -\frac{\partial H}{\partial x}(x, p) \right)$  the *Hamiltonian vector field* associated to  $H$ . This has a more intrinsic (geometric) definition in terms of the canonical symplectic structure  $\omega$  on  $T^*M$ . In fact,  $X_H$  is the unique vector field that satisfies

$$\omega(X_H(x, p), \cdot) = d_x H(\cdot) \quad \forall (x, p) \in T^*M.$$

For this reason, it is sometime called *symplectic gradient of  $H$* . It is easy to check from both definitions that - only in the autonomous case - the Hamiltonian is a *prime integral of the motion*, i.e., it is constant along the solutions of these equations.

Now, we would like to explain what is the relation between the Euler-Lagrange flow and the Hamiltonian one. It follows easily from the definition of Hamiltonian (and Fenchel transform) that for each  $(x, v) \in TM$  and  $(x, p) \in T^*M$  the following inequality holds:

$$(3) \quad \langle p, v \rangle_x \leq L(x, v) + H(x, p).$$

This is called *Fenchel inequality* (or *Fenchel-Legendre inequality*) and plays a crucial role in the study of Lagrangian and Hamiltonian dynamics and in the variational methods that we are going to describe. In particular, equality holds if and only if  $p = \partial L / \partial v(x, v)$ . One can therefore introduce the following diffeomorphism between  $TM$  and  $T^*M$ , known as *Legendre transform*:

$$(4) \quad \begin{aligned} \mathcal{L} : TM &\longrightarrow T^*M \\ (x, v) &\longmapsto \left( x, \frac{\partial L}{\partial v}(x, v) \right). \end{aligned}$$

Moreover the following relation with the Hamiltonian holds:

$$H \circ \mathcal{L}(x, v) = \left\langle \frac{\partial L}{\partial v}(x, v), v \right\rangle_x - L(x, v).$$

A crucial observation is that this diffeomorphism  $\mathcal{L}$  represents a conjugation between the two flows, namely the Euler-Lagrange flow on  $TM$  and the Hamiltonian flow on  $T^*M$ ; in other words, the following diagram commutes:

$$\begin{array}{ccc} TM & \xrightarrow{\Phi_t^L} & TM \\ \mathcal{L} \downarrow & & \downarrow \mathcal{L} \\ T^*M & \xrightarrow{\Phi_t^H} & T^*M \end{array}$$

**Remark 2.4.** Since  $\mathcal{L}$  and the Hamiltonian flow  $\Phi^H$  are both  $C^1$ , then it follows from the commutative diagram above that the Euler-Lagrange flow is also  $C^1$ .

Therefore one can equivalently study the Euler-Lagrange flow or the Hamiltonian flow, obtaining in both cases information on the dynamics of the system. Each of these equivalent approaches will provide different tools and advantages, which may prove very useful to understand the dynamical properties of the system. For instance, the tangent space is the natural setting for the classical calculus of variations and for Mather and Mañé's approaches (sections 4 and 5); on the other hand, the cotangent space is equipped with a canonical symplectic structure, which allows one to use several symplectic topological tools, coming from the study of Lagrangian graphs, Hofer's theory, Floer Homology, *etc* ... Moreover, a particular fruitful approach in  $T^*M$  is the so-called *Hamilton-Jacobi method* (or *Weak KAM theory*), which is concerned with the study of *solutions* and *subsolutions* of Hamilton-Jacobi equations. In a certain sense, this approach represents the functional analytical counterpart of the above-mentioned variational approach (section 6). In the following sections we shall provide a complete description of these methods and their implications to the study of the dynamics of the system.

### 3. A CARTOON EXAMPLE: PROPERTIES OF MEASURES AND ORBITS ON KAM TORI

Before entering into the details of Mather's work, we would like to discuss some properties of orbits and invariant probability measures that are supported on *KAM tori*. This will provide us with a better understanding of the ideas behind Mather's theory and allow to see in what sense these *action-minimizing sets* - namely, what we shall call *Mather*, *Aubry* and *Mañé sets* - represent a generalization of KAM tori.

Let  $H : T^*\mathbb{T}^d \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $L : T\mathbb{T}^d \rightarrow \mathbb{R}$  its associated Tonelli Lagrangian and let us denote by  $\Phi^H$  and  $\Phi^L$  the respective flows. Observe that one can identify  $T^*\mathbb{T}^d$  and  $T\mathbb{T}^d$  with  $\mathbb{T}^d \times \mathbb{R}^d$ . First of all, let us define what we mean by KAM torus.

**Definition 3.1 (KAM Torus).**  $\mathcal{T} \subset \mathbb{T}^d \times \mathbb{R}^d$  is a (maximal) KAM torus with rotation vector  $\rho$  if:

- i)  $\mathcal{T} \subset \mathbb{T}^d \times \mathbb{R}^d$  is a  $C^1$  Lagrangian graph, *i.e.*,  $\mathcal{T} = \{(x, c + du) : x \in \mathbb{T}^d\}$ , where  $c \in \mathbb{R}^d$  and  $u : \mathbb{T}^d \rightarrow \mathbb{R}$ ;
- ii)  $\mathcal{T}$  is invariant under the Hamiltonian flow  $\Phi_t^H$  generated by  $H$ ;
- iii) the Hamiltonian flow on  $\mathcal{T}$  is conjugated to a uniform rotation on  $\mathbb{T}^d$ ; *i.e.*, there exists a diffeomorphism  $\varphi : \mathbb{T}^d \rightarrow \mathcal{T}$  such that  $\varphi^{-1} \circ \Phi_t^H \circ \varphi = R_\rho^t$ ,  $\forall t \in \mathbb{R}$ , where  $R_\rho^t : x \rightarrow x + \rho t \pmod{\mathbb{Z}^d}$ . In other words, the following diagram commutes:

$$\begin{array}{ccc} \mathcal{T} & \xrightarrow{\Phi_t^H} & \mathcal{T} \\ \uparrow \varphi & & \uparrow \varphi \\ \mathbb{T}^d & \xrightarrow{R_\rho^t} & \mathbb{T}^d \end{array}$$

**Remark 3.2.** (i) First of all, let us remark that, up to some small technicalities in the proofs, most of the results that we shall discuss in this section, will continue to hold if we assume  $\mathcal{T}$  to be only Lipschitz rather than  $C^1$ .

(ii) If  $\rho$  is rationally independent, *i.e.*,  $\rho \cdot k \neq 0$  for all  $k \in \mathbb{Z}^d \setminus \{0\}$ , then each orbit is dense on it (*i.e.*, the motion is said to be *minimal*). Therefore, it supports a unique invariant probability measure. Consider the case in which  $\rho$  admits some resonance, *i.e.*, there exists  $k \in \mathbb{Z}^d \setminus \{0\}$  such that  $\rho \cdot k = 0$ . The set of resonances



of  $\rho$  forms a module over  $\mathbb{Z}$ . Let  $r$  denote its rank. Then,  $\mathcal{T}$  is foliated by an  $r$ -dimensional family of  $(d-r)$ -tori, on each of which the motion is minimal.

(iii) If  $\rho$  is rationally independent then a classical result by Michel Herman implies that this torus is automatically Lagrangian (see [29, Proposition 3.2]).

(iv) Since  $\mathcal{T}$  is Lagrangian and invariant under the Hamiltonian flow, then the Hamiltonian  $H$  is constant on it, *i.e.*,  $H(x, c + du) = E_c$  for some  $E_c \in \mathbb{R}$  (it follows from the definition of the Hamiltonian flow). In particular,  $u$  is a classical solution of Hamilton-Jacobi equation  $H(x, c + dv) = k$ . It is easy to check that, for a fixed  $c$ , all solutions have the same energy. In fact, let  $u$  and  $v$  be two solutions. Then, the function  $u - v$  will have at least a critical point  $x_0$  and at  $x_0$  the two differentials must coincide:  $d_{x_0}u = d_{x_0}v$ . This implies that  $H(x_0, c + d_{x_0}u) = H(x_0, c + d_{x_0}v)$ . Moreover,  $E_c$  is the least possible  $k \in \mathbb{R}$  such that there can exist subsolutions of the equation  $H(x, c + dv) = k$ , namely  $v \in C^1(\mathbb{T}^d)$  satisfying  $H(x, c + dv) \leq k$  for all  $x \in \mathbb{T}^d$ . The proof is essentially the same as above. In fact, since the function  $u - v$  has at least a critical point  $x_0$  and at  $x_0$  the two differentials coincide, then  $E_c = H(x_0, c + d_{x_0}u) = H(x_0, c + d_{x_0}v) \leq k$ .

Let us start by studying the properties of invariant probability measures that are supported on  $\mathcal{T}$ . Let  $\mu^*$  be an ergodic invariant probability measure supported on  $\mathcal{T}$ . It will be more convenient at this point to work in the Lagrangian setting. Let us consider the correspondent invariant probability measure for the Euler-Lagrange flow  $\Phi^L$ , obtained using the Legendre transform:  $\mu = \mathcal{L}^*\mu^*$ , where  $\mathcal{L}^*$  denotes the pull-back. We would like to point out some properties of this measure and see how  $\mathcal{T}$  can be characterized using measures satisfying such properties.

### Properties:

- (1) First of all, observe that  $\int v d\mu = \rho$ . To see this, let us consider the universal cover of  $\mathbb{T}^d$ , *i.e.*,  $\mathbb{R}^d$ , and denote by:
  - $\tilde{\mathcal{T}} := \{(q, c + d\tilde{u}(q)), q \in \mathbb{R}^d\}$  the lift of  $\mathcal{T}$  to  $\mathbb{R}^d$ , where  $\tilde{u}$  is a periodic extension of  $u$ ;
  - $\Phi_t^H$  the suspension of the Hamiltonian flow  $\Phi_t^H$  to  $\mathbb{R}^d \times \mathbb{R}^d$ ;
  - $\tilde{\varphi}$  the lift of  $\varphi$ , *i.e.*,  $\varphi : \mathbb{R}^d \rightarrow \tilde{\mathcal{T}}$  such that the following diagram commutes:

$$\begin{array}{ccc} \tilde{\mathcal{T}} & \xrightarrow{\Phi_t^H} & \tilde{\mathcal{T}} \\ \tilde{\varphi} \uparrow & & \uparrow \tilde{\varphi} \\ \mathbb{R}^d & \xrightarrow{\tilde{R}_\rho^t} & \mathbb{R}^d \end{array}$$

where  $\tilde{R}_\rho^t : q \rightarrow q + \rho t$ . Moreover,  $\tilde{\varphi}$  has the form  $\tilde{\varphi}(q) = (\tilde{\varphi}_q(q), c + d\tilde{u}(\tilde{\varphi}_q(q)))$  and  $\tilde{\varphi}_q$  satisfies  $\tilde{\varphi}_q(q + k) = \tilde{\varphi}_q(q) + k$  for each  $k \in \mathbb{Z}^d$ .

Now, let  $(x_0, c + du(x_0))$  be a point in the support of  $\mu^*$  and consider the corresponding orbit  $\gamma_{x_0}(t) := \pi \Phi_t^H(x_0, c + du(x_0))$ , where  $\pi : \mathbb{T}^*\mathbb{T}^d \rightarrow \mathbb{T}^d$  is the canonical projection. Denote by  $\tilde{\gamma}_{x_0}(t)$  its lift to  $\mathbb{R}^d$ . Then, using the ergodic theorem (remember that we are assuming that  $\mu^*$ , and hence  $\mu$ , is ergodic), we have that for a generic point  $(x_0, c + du(x_0))$  in the support of

$\mu^*$  the following holds:

$$\begin{aligned}
\int_{\mathbb{T}^d} v d\mu &= \lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^n \dot{\gamma}_{x_0}(t) dt = \lim_{n \rightarrow +\infty} \frac{\widetilde{\gamma}_{x_0}(n) - \widetilde{\gamma}_{x_0}(0)}{n} = \\
&= \lim_{n \rightarrow +\infty} \frac{\widetilde{\gamma}_{x_0}(n)}{n} = \lim_{n \rightarrow +\infty} \frac{\widetilde{\varphi}_q(\widetilde{\varphi}_q^{-1}(\widetilde{\gamma}_{x_0}(0)) + n\rho)}{n} = \\
&= \lim_{n \rightarrow +\infty} \left[ \frac{\widetilde{\varphi}_q(\widetilde{\varphi}_q^{-1}(\widetilde{\gamma}_{x_0}(0)) + \{n\rho\})}{n} + \frac{[n\rho]}{n} \right] = \\
&= \rho
\end{aligned}$$

where  $[\cdot]$  and  $\{\cdot\}$  denote respectively the integer and fractional part of the vector (component by component).

**Remark 3.3.** We shall say that  $\mu$  has *rotation vector*  $\rho$ . See also Section 4.

- (2) Let  $f : \mathbb{T}^d \rightarrow \mathbb{R}$  be a  $C^1$  function. Then,  $\int df(x) \cdot v d\mu = 0$ . In fact, take  $(x_0, v_0) \in \text{supp } \mu$  and denote the corresponding orbit by  $(x_t, v_t) = \Phi_t^L(x_0, v_0)$ . Since  $\mu$  is invariant, then  $\Phi_{t*}^L \mu = \mu$  and therefore:

$$\begin{aligned}
\int_{\mathbb{T}^d} df(x) \cdot v d\mu &= \frac{1}{T} \int_0^T dt \int_{\mathbb{T}^d} df(x) \cdot v d\Phi_{t*}^L \mu = \\
&= \frac{1}{T} \int_0^T dt \int_{\mathbb{T}^d} df(x_t) \cdot v_t d\mu = \\
&= \frac{1}{T} \int_{\mathbb{T}^d} d\mu \int_0^T df(x_t) \cdot v_t dt = \\
&= \int_{\mathbb{T}^d} \frac{f(x_T) - f(x_0)}{T} d\mu \xrightarrow{T \rightarrow +\infty} 0,
\end{aligned}$$

since  $f$  is bounded. A measure satisfying this condition is said to be *closed* (see also Addendum 4.A). Observe that in the above proof we have not used anything else than the invariance of  $\mu$ . See also Proposition 4.7.

- (3) Recall that  $E_c$  is the *energy* of  $\mathcal{T}$ , i.e.,  $H(x, c + du) = E_c$ . We want to show that this energy value is somehow related to the average value of  $L$  on the support of  $\mu$ . Observe in fact that:

$$\int_{\mathbb{T}^d} L(x, v) d\mu = \int_{\mathbb{T}^d} [L(x, v) - (c + du) \cdot v] d\mu + \int_{\mathbb{T}^d} (c + du) \cdot v d\mu.$$

Using the property just pointed out in (2) and observing that, because of the Fenchel-Legendre (in)equality (3), along the orbits  $(x_t, v_t)$  in the support of  $\mu$  we have  $(c + du(x_t)) \cdot v_t = L(x_t, v_t) + H(x, c + du(x_t)) = L(x_t, v_t) + E_c$ , we can conclude that:

$$\begin{aligned}
\int_{\mathbb{T}^d} L(x, v) d\mu &= - \int_{\mathbb{T}^d} E_c d\mu + \int_{\mathbb{T}^d} c \cdot v d\mu = \\
&= -E_c + c \cdot \int_{\mathbb{T}^d} v d\mu = -E_c + c \cdot \rho.
\end{aligned}$$

In particular,  $\int_{\mathbb{T}^d} [L(x, v) - c \cdot v] d\mu = -E_c$ . Observe that this new Lagrangian  $L_c(x, v) = L(x, v) - c \cdot v$  is still Tonelli and it is immediate to check that it has the same Euler-Lagrange flow as  $L$ .

We can now show this first result.

**Proposition 3.4.** *If  $\tilde{\mu}$  is another invariant probability measure of  $\Phi^L$ , then*

$$(5) \quad \int_{\mathbb{T}^d} L_c(x, v) d\tilde{\mu} \geq \int_{\mathbb{T}^d} L_c(x, v) d\mu.$$

Therefore,  $\mu$  *minimizes* the average value of  $L_c$  (or *action* of  $L_c$ ) among all invariant probability measures of  $\Phi_t^{L_c}$  (or  $\Phi_t^L$  since they are the same).

WARNING: in general, it does not minimize the action of  $L$ !

**Proof.** The proof is an easy application of Fenchel-Legendre inequality (3). Observe that in the support of  $\tilde{\mu}$ , differently from what happens on the support of  $\mu$  (see property (3) above), this is not an equality, *i.e.*,  $(c + du(x)) \cdot v \leq L(x, v) + H(x, c + du(x)) = L(x, v) + E_c$ . Then:

$$\begin{aligned} \int_{\mathbb{T}^d} L_c(x, v) d\tilde{\mu} &= \int_{\mathbb{T}^d} [L(x, v) - (c + du) \cdot v] d\tilde{\mu} \geq \\ &\geq - \int_{\mathbb{T}^d} E_c d\tilde{\mu} = -E_c = \int_{\mathbb{T}^d} L_c(x, v) d\mu. \end{aligned}$$

□

**Remark 3.5.** (i) It follows from the above proposition that

$$E_c = - \min \left\{ \int_{\mathbb{T}^d} L_c(x, v) d\tilde{\mu} : \tilde{\mu} \text{ is a } \Phi^L\text{-invariant probability measure} \right\}.$$

The minimizing action on the right-hand side is also denoted by  $\alpha(c)$  (see the definition of Mather's  $\alpha$ -function in Section 4).

(ii) A measure  $\mu$  that satisfies the inequality (5) for all invariant probability measures  $\tilde{\mu}$ , is called *c-action minimizing measure* (or *Mather's measure with cohomology class c*); see Section 4.

(iii) It is easy to see from the proof of Proposition 3.4 and Fenchel-Legendre inequality, that if  $\text{supp } \tilde{\mu}$  is not contained in  $\mathcal{L}^{-1}(\mathcal{T})$ , then the inequality in (5) is strict.

(iii) In particular, using the previous remark and (ii) in Remark 3.2, we obtain that

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{ \text{supp } \mu : \mu \text{ is a c-action minimizing measure} \}.$$

The set on the right-hand side is often denoted by  $\widetilde{\mathcal{M}}_c$  and called *Mather set of cohomology class c*. It will be defined in Section 4.

Let us now observe that although  $\mu$  does not minimize the action of  $L$  among all invariant probability measures (we have pointed out in fact that it minimizes the action of a modified Lagrangian, with the same Euler Lagrange flow as  $L$ ), then it does minimize it if we put some extra constraints.

**Proposition 3.6.** *If  $\tilde{\mu}$  is another invariant probability measure of  $\Phi^L$  with rotation vector  $\rho$  (in the sense of property (1) above), then*

$$(6) \quad \int_{\mathbb{T}^d} L(x, v) d\tilde{\mu} \geq \int_{\mathbb{T}^d} L(x, v) d\mu.$$

**Proof.** The proof is the same as before, using Fenchel-Legendre inequality (3), the fact that  $\int v d\tilde{\mu} = \rho$  and property (3) above. In fact:

$$\begin{aligned} \int_{\mathbb{T}^d} L(x, v) d\tilde{\mu} &= \int_{\mathbb{T}^d} [L(x, v) - (c + du) \cdot v] d\tilde{\mu} + \int_{\mathbb{T}^d} (c + du) \cdot v d\tilde{\mu} \geq \\ &\geq -E_c + c \cdot \int_{\mathbb{T}^d} v d\tilde{\mu} = \\ &= -E_c + c \cdot \rho = \int_{\mathbb{T}^d} L(x, v) d\mu. \end{aligned}$$

□

**Remark 3.7.** (i) It follows from the above proposition that

$$-E_c + c \cdot \rho = \min \left\{ \int_{\mathbb{T}^d} L(x, v) d\tilde{\mu} : \tilde{\mu} \text{ is a } \Phi^L\text{-inv. prob. meas. with rot. vector } \rho \right\}.$$

The minimizing action on the right-hand side is also denoted by  $\beta(\rho)$  (see the definition of Mather's  $\beta$ -function in Section 4).

(ii) A measure  $\mu$  that satisfies the inequality (6) for all invariant probability measures  $\tilde{\mu}$  with rotation vector  $\rho$ , is called *action minimizing measure* (or *Mather's measure*) *with rotation vector*  $\rho$ ; see Section 4.

(iii) It is easy to see from the proof of Proposition 3.6 and Fenchel-Legendre inequality, that if  $\tilde{\mu}$  has rotation vector  $\rho$ , but its support is not contained in  $\mathcal{L}^{-1}(\mathcal{T})$ , then the inequality in (6) is strict.

(iv) In particular, using the previous remark and (ii) in Remark 3.2, we obtain that

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{ \text{supp } \mu : \mu \text{ is an action minimizing measure with rotation vector } \rho \}.$$

The set on the right-hand side is often denoted by  $\widetilde{\mathcal{M}}^\rho$  and called *Mather set of homology class*  $\rho$ . It will be defined in Section 4.

Summarizing, if  $\mu$  is any invariant probability measure of  $\Phi^L$  supported on  $\mathcal{L}^{-1}(\mathcal{T})$ , then:

- $\mu$  minimizes the *action* of  $L(x, v) - c \cdot v$  amongst all invariant probability measures of  $\Phi^L$ ;
- $\mu$  minimizes the *action* of  $L(x, v)$  amongst all invariant probability measures of  $\Phi^L$  with rotation vector  $\rho$ .

Let us now shift our attention to orbits on KAM tori and see what are the properties that they enjoy. Let  $(x_0, c + d_{x_0}u) \in \mathcal{T}$  a point on the KAM torus and consider its orbit under the Hamiltonian flow, *i.e.*,  $\gamma(t) = \pi \Phi_t^H(x_0, c + d_{x_0}u)$ , where as usual  $\pi : \mathbb{T}^* \mathbb{T}^d \rightarrow \mathbb{T}^d$  denotes the canonical projection along the fiber. Let us fix any times  $a < b$  and consider the corresponding *Lagrangian action* of this curve. Using Fenchel-Legendre (in)equality we get:

$$\begin{aligned} \int_a^b L(\gamma(t), \dot{\gamma}(t)) dt &= \int_a^b \left( (c + d_{\gamma(t)}u) \dot{\gamma}(t) - H(\gamma(t), c + d_{\gamma(t)}u) \right) dt = \\ &= \int_a^b c \dot{\gamma}(t) dt + u(\gamma(b)) - u(\gamma(a)) - E_c(b - a). \end{aligned}$$

Therefore, considering as above the action of the modified Lagrangian  $L_c(x, v)$ , we get:

$$(7) \quad \int_a^b L_c(\gamma(t), \dot{\gamma}(t)) dt = u(\gamma(b)) - u(\gamma(a)) - E_c(b - a).$$

Let us now take any other absolutely continuous curve  $\xi : [a, b] \rightarrow \mathbb{T}^d$  with the same endpoints as  $\gamma$ , i.e.,  $\xi(a) = \gamma(a)$  and  $\xi(b) = \gamma(b)$ . Proceeding as before and using Fenchel-Legendre inequality, we obtain:

$$\begin{aligned} \int_a^b L(\xi(t), \dot{\xi}(t)) dt &\geq \int_a^b \left( (c + d_{\xi(t)}u) \dot{\xi}(t) - H(\xi(t), c + d_{\xi(t)}u) \right) dt = \\ &= \int_a^b c \dot{\xi}(t) dt + u(\xi(b)) - u(\xi(a)) - E_c(b - a). \end{aligned}$$

Hence, using (7) and the fact that  $\xi(a) = \gamma(a)$  and  $\xi(b) = \gamma(b)$ , we can conclude that:

$$(8) \quad \int_a^b L_c(\gamma(t), \dot{\gamma}(t)) dt \leq \int_a^b L_c(\xi(t), \dot{\xi}(t)) dt.$$

Therefore for any times  $a < b$ ,  $\gamma$  is the curve that minimizes the action of  $L_c$  over all absolutely continuous curves  $\xi : [a, b] \rightarrow \mathbb{T}^d$  with  $\xi(a) = \gamma(a)$  and  $\xi(b) = \gamma(b)$ . Actually something more is true. Let us consider a curve with the same endpoints, but a different time-length, i.e.,  $\xi : [a', b'] \rightarrow \mathbb{T}^d$  with  $a' < b'$  and such that  $\xi(a') = \gamma(a)$  and  $\xi(b') = \gamma(b)$ . Proceeding as above, one obtains

$$\int_{a'}^{b'} L_c(\xi(t), \dot{\xi}(t)) dt \geq u(\xi(b')) - u(\xi(a')) - E_c(b' - a')$$

and consequently

$$\int_a^b \left( L_c(\gamma(t), \dot{\gamma}(t)) + E_c \right) dt \leq \int_{a'}^{b'} \left( L_c(\xi(t), \dot{\xi}(t)) + E_c \right) dt.$$

Hence, for any times  $a < b$ ,  $\gamma$  minimizes the action of  $L_c + E_c$  amongst all absolutely continuous curves  $\xi$  that connect  $\gamma(a)$  to  $\xi(b) = \gamma(b)$  in any given time (adding a constant does not change the Euler-Lagrange flow).

We have just proved the following proposition.

**Proposition 3.8.** *For any given  $a < b$ , the projection  $\gamma$  of any orbit on  $\mathcal{T}$  minimizes the action of  $L_c$  amongst all absolutely continuous curves that connect  $\gamma(a)$  to  $\gamma(b)$  in time  $b - a$ . Furthermore,  $\gamma$  minimizes the action of  $L_c + E_c$  amongst all absolutely continuous curves that connect  $\gamma(a)$  to  $\gamma(b)$  in any given time length.*

**Remark 3.9.** (i) A curve  $\gamma : \mathbb{R} \rightarrow \mathbb{T}^d$  such that for any  $a < b$ ,  $\gamma|_{[a, b]}$  minimizes the action of  $L_c$  amongst all absolutely continuous curves that connect  $\gamma(a)$  to  $\gamma(b)$  in time  $b - a$ , is called a *c-global minimizer* of  $L$ .

(ii) A curve  $\gamma : \mathbb{R} \rightarrow \mathbb{T}^d$  such that for any  $a < b$ ,  $\gamma|_{[a, b]}$  minimizes the action of  $L_c$  amongst all absolutely continuous curves that connect  $\gamma(a)$  to  $\gamma(b)$  without any restriction on the time length, is called a *c-time free minimizer* of  $L$ .

(iii) Proposition 3.8 can be restated by saying that the projection of each orbit on  $\mathcal{T}$  is a *c-global minimizer* of  $L$  and a *c-time free minimizer* of  $L + E_c$ . Moreover, it follows easily from the proof and Fenchel-Legendre inequality, that if the curve  $\xi$  does not lie on  $\mathcal{T}$ , then the inequality in (8) is strict.

In particular, we obtain that:

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a c-global minimizer of } L \text{ and } t \in \mathbb{R}\}.$$

The set on the right-hand side is often denoted by  $\tilde{\mathcal{N}}_c$  and called *Mañé set of cohomology class  $c$* . It will be defined in Section 5.

Actually, these curves on  $\mathcal{T}$  are more than  $c$ -global minimizers. As we shall see, they satisfy a more restrictive condition. In order to introduce it, let us introduce what is called the *Mañé potential*. Let  $x_1, x_2 \in \mathbb{T}^d$  and let us denote by  $h_c^T(x_1, x_2)$  the minimal action of  $L_c$  along curves that connect  $x_1$  to  $x_2$  in time  $T$ . We want to consider the infimum of these quantities for all positive times. Of course, without any “correction” this infimum might be  $-\infty$ . Hence, let us consider the action of  $L_c + k$  for some  $k \in \mathbb{R}$  and define:

$$(9) \quad \phi_{c,k}(x_1, x_2) := \inf_{T>0} (h_c^T(x_1, x_2) + kT).$$

We want to see for which values of  $k$  this quantity is well-defined, *i.e.*, it is  $> -\infty$ . Let  $\gamma : [0, T] \rightarrow \mathbb{T}^d$  be any absolutely continuous curve such that  $\gamma(0) = x_1$  and  $\gamma(T) = x_2$ . Then, using Fenchel Legendre inequality:

$$\begin{aligned} \int_0^T L_c(\gamma(t), \dot{\gamma}(t)) dt + kT &\geq \int_0^T (du(\gamma(t)) \cdot \dot{\gamma}(t) - H(\gamma(t), c + d_{\gamma(t)}u)) dt + kT = \\ &= u(\gamma(T)) - u(\gamma(0)) + (k - E_c)T = \\ (10) \quad &= u(x_2) - u(x_1) + (k - E_c)T. \end{aligned}$$

Taking the infimum over all curves connecting  $x_1$  to  $x_2$ , we obtain:

$$\phi_{c,k}(x_1, x_2) \geq u(x_2) - u(x_1) + (k - E_c)T.$$

Therefore, if  $k \geq E_c$ , then  $\phi_{c,k}(x_1, x_2) > -\infty$  for all  $x_1, x_2 \in \mathbb{T}^d$ . On the other hand, it is quite easy to show that for  $k < E_c$ ,  $\phi_{c,k}(x_1, x_2) = -\infty$  for all  $x_1, x_2 \in \mathbb{T}^d$ . Let us observe the following:

- If  $k < E_c$ , then  $\phi_{c,k}(x, x) = -\infty$  for all  $x \in \mathbb{T}^d$ . In fact, we know that the orbit  $\gamma(t) = \pi \Phi_t^H(x, c + du(x))$  is recurrent, *i.e.*, there exist  $T_n \rightarrow +\infty$  such that  $\delta_n := \|\gamma(T_n) - x\| \rightarrow 0$  as  $n \rightarrow \infty$ . Consider the curve  $\xi : [0, T_n + \delta_n] \rightarrow \mathbb{T}^d$  obtained by joining the curve  $\gamma$  to the unit speed geodesic connecting  $\gamma(T_n)$  to  $x$ . Then, proceeding as before and using Fenchel-Legendre inequality and the continuity of  $u$ , we obtain:

$$\begin{aligned} \phi_{c,k}(x, x) &\leq \int_0^{T_n} L_c(\gamma, \dot{\gamma}) dt + \int_{T_n}^{T_n + \delta_n} L_c(\xi, \dot{\xi}) dt + k(T_n + \delta_n) \leq \\ &\leq u(\gamma(T_n)) - u(x) + A\delta_n + (k - E_c)T_n + k\delta_n \xrightarrow{n \rightarrow +\infty} -\infty, \end{aligned}$$

where  $A := \max_{\|v\|=1} |L_c(x, v)|$ .

- It is also easy to check that for any  $k$ ,  $\phi_{c,k}$  satisfies a sort of triangular inequality:

$$\phi_{c,k}(x, y) \leq \phi_{c,k}(x, z) + \phi_{c,k}(z, y) \quad \forall x, y, z \in \mathbb{T}^d.$$

For the proof of this, observe that any absolutely continuous curve connecting  $x$  to  $z$  and any absolutely continuous curve connecting  $z$  to  $y$ , determine a curve from  $x$  to  $y$ . Passing to the infimum, one gets the result.

Our claim follows easily from these two observations. In fact, let  $k < E_c$  and  $x_1, x_2 \in \mathbb{T}^d$ . Then:

$$\phi_{c,k}(x_1, x_2) \leq \phi_{c,k}(x_1, x_1) + \phi_{c,k}(x_1, x_2) = -\infty.$$

Summarizing, we have proved the following proposition.

**Proposition 3.10.**

$$\begin{aligned}
E_c &= \inf\{k \in \mathbb{R} : \phi_{c,k}(x, y) > -\infty \text{ for all } x, y \in \mathbb{T}^d\} = \\
&= \sup\{k \in \mathbb{R} : \phi_{c,k}(x, y) = -\infty \text{ for all } x, y \in \mathbb{T}^d\} = \\
&= \sup\{k \in \mathbb{R} : \phi_{c,k}(x, x) = -\infty \text{ for all } x \in \mathbb{T}^d\} = \\
&= \sup\{k \in \mathbb{R} : \phi_{c,k}(x, x) = -\infty \text{ for some } x \in \mathbb{T}^d\}.
\end{aligned}$$

Observe that the second and fourth equality are simply a consequence of the triangular inequality.

Let us go back to our  $c$ -global minimizing curves. We have seen (Proposition 3.8) that if  $\gamma$  is the projection of an orbit on  $\mathcal{T}$ , then it is a  $c$ -global minimizer of  $L$  and furthermore it is a  $c$ -time free minimizer of  $L + E_c$ . The last property can be rewritten in terms of the Mañé potential:

$$\int_a^b (L_c(\gamma(t), \dot{\gamma}(t)) + E_c) dt = \phi_{c,E_c}(\gamma(a), \gamma(b)) \quad \text{for any } a < b.$$

Therefore these curves seems to be related to the Mañé potential corresponding to  $k = E_c$  (the least value for which it is defined). Observe that in this case, we obtain from (10) that  $\phi_{c,E_c}(x, y) \geq u(y) - u(x)$  for all  $x, y \in \mathbb{T}^d$  and consequently:

$$(11) \quad \phi_{c,E_c}(y, x) \geq u(x) - u(y) = -(u(y) - u(x)) \geq -\phi_{c,E_c}(x, y),$$

which provides a lower bound on the action needed to “go back” from  $y$  to  $x$ . In particular, if  $\xi : [a, b] \rightarrow \mathbb{T}^d$  is an absolutely continuous curve, then:

$$\int_a^b (L_c(\xi, \dot{\xi}) + E_c) dt \geq \phi_{c,E_c}(\xi(a), \xi(b)) \geq -\phi_{c,E_c}(\xi(b), \xi(a)).$$

Question: do there exist curves for which these inequalities are equalities. Observe, that if such a curve exists, then it must necessarily be a  $c$ -minimizer. We shall show that the answer is affirmative and characterize such curves in our case.

**Proposition 3.11.** *Let  $\gamma$  be the projection of an orbit on  $\mathcal{T}$ . Then, for each  $a < b$  we have:*

$$\int_a^b (L_c(\gamma, \dot{\gamma}) + E_c) dt = \phi_{c,E_c}(\gamma(a), \gamma(b)) = -\phi_{c,E_c}(\gamma(b), \gamma(a)).$$

**Remark 3.12.** (i) A curve  $\gamma : \mathbb{R} \rightarrow \mathbb{T}^d$  satisfying the conditions in Proposition 3.11 is called *c-regular global minimizer* or *c-static curve* of  $L$ . Observe that the adjective *regular* (coined by John Mather) has no relation to the smoothness of the curve, since this curve will be as smooth as all other solutions of the Euler-Lagrange flow (depending on the regularity of the Lagrangian).

(ii) Since a  $c$ -regular global minimizer is a  $c$ -global minimizer (see the comment above), it follows from (iii) in Remark 3.9 that these are all and only the  $c$ -regular global minimizers of  $L$ . Therefore:

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a } c\text{-regular global minimizer of } L \text{ and } t \in \mathbb{R}\}.$$

The set on the right-hand side is often denoted by  $\tilde{\mathcal{A}}_c$  and called *Aubry set of cohomology class  $c$* . It will be defined in Section 5.

**Proof. [Proposition 3.11]** Let  $\gamma(t) := \pi \Phi_t^H(x, c + du(x))$  and consider  $a < b$ . We already know from Proposition 3.8 that

$$\int_a^b (L_c(\gamma, \dot{\gamma}) + E_c) dt = \phi_{c,E_c}(\gamma(a), \gamma(b)).$$

Since the orbit is recurrent, there exist  $T_n \rightarrow +\infty$  such that  $\delta_n := \|\gamma(b + T_n) - \gamma(a)\| \rightarrow 0$ . Consider the curve  $\xi : [b, b + T_n + \delta_n] \rightarrow \mathbb{T}^d$  obtained by joining the curve  $\gamma|_{[b, b + T_n]}$  to the unit speed geodesic connecting  $\gamma(b + T_n)$  to  $\gamma(a)$ . Then, using Fenchel-Legendre inequality and the continuity of  $u$ , we obtain:

$$\begin{aligned} \phi_{c, E_c}(\gamma(b), \gamma(a)) &\leq \int_b^{b+T_n} L_c(\gamma, \dot{\gamma}) dt + \int_{b+T_n}^{b+T_n+\delta_n} L_c(\xi, \dot{\xi}) dt + E_c(T_n + \delta_n) \leq \\ &\leq u(\gamma(b + T_n)) - u(\gamma(b)) + (A + E_c)\delta_n \xrightarrow{n \rightarrow +\infty} u(\gamma(a)) - u(\gamma(b)), \end{aligned}$$

where  $A := \max_{\|v\|=1} |L_c(x, v)|$ . Therefore, using (10), we obtain:

$$\begin{aligned} \phi_{c, E_c}(\gamma(b), \gamma(a)) &\leq u(\gamma(a)) - u(\gamma(b)) = -(u(\gamma(b)) - u(\gamma(a))) = \\ &= -\phi_{c, E_c}(\gamma(a), \gamma(b)) \end{aligned}$$

that, together with (11), allows us to deduce that  $\phi_{c, E_c}(\gamma(b), \gamma(a)) = -\phi_{c, E_c}(\gamma(a), \gamma(b))$  and conclude the proof.  $\square$

#### 4. ACTION-MINIMIZING MEASURES: MATHER SETS

In this section, inspired by what happens for invariant measures and orbits on KAM tori, we would like to attempt a similar approach for general Tonelli Lagrangians on compact manifolds, and prove the existence of some analogous compact invariant subsets for the Euler-Lagrange (or Hamiltonian) flow. In many senses, these sets will resemble and generalize KAM tori, even when these do not exist or it does not make sense to speak of them.

Let us start by studying invariant probability measures of the system and their action-minimizing properties. Then, we shall use them to define a first family of invariant sets: the *Mather sets* (see Remark 3.5 (iii) and Remark 3.7 (iii)). Let  $\mathfrak{M}(L)$  be the space of probability measures  $\mu$  on  $TM$  that are invariant under the Euler-Lagrange flow of  $L$  and such that  $\int_{TM} L d\mu < \infty$  (finite action). It is easy to see in the case of an autonomous Tonelli Lagrangian, that this set is non-empty. In fact, recall that because of the conservation of the energy  $E(x, v) := H \circ \mathcal{L}(x, v) = \langle \frac{\partial L}{\partial v}(x, v), v \rangle_x - L(x, v)$  along the motions, each energy level of  $E$  is compact (it follows from the superlinearity condition) and invariant under  $\Phi_t^L$ . It is a well-known result by Kryloff and Bogoliouboff [31] that a flow on a compact metric space has at least an invariant probability measure.

**Proposition 4.1.** *Each non-empty energy level  $\mathcal{E}(E) := \{E(x, v) = E\}$  contains at least one invariant probability measure of  $\Phi_t^L$ .*

**Proof.** Let  $(x_0, v_0) \in \mathcal{E}(E)$  and consider the curve  $\gamma(t) := \pi \Phi_t^L(x_0, v_0)$ , where  $\pi : TM \rightarrow M$  denotes the canonical projection. For each  $T > 0$ , we can consider the probability measure  $\mu_T$  uniformly distributed on the piece of curve  $(\gamma(t), \dot{\gamma}(t))$  for  $t \in [0, T]$ , i.e.,

$$\int f d\mu_T = \frac{1}{T} \int_0^T f(\gamma(t), \dot{\gamma}(t)) dt \quad \forall f \in C(\mathcal{E}(E)).$$

The family  $\{\mu_T\}_T$  is precompact in the weak\* topology; in fact the set of probability measures on  $\mathcal{E}(E)$  is contained in the closed unit ball in the space of signed Borel measures of  $\mathcal{E}(E)$ , which is weak\* compact by Banach-Alaouglu theorem. Therefore, I can extract a converging subsequence  $\{\mu_{T_n}\}_n$ , such that  $\mu_{T_n} \rightarrow \mu$  in



the weak\* topology as  $T_n \rightarrow \infty$ . We want to prove that  $\mu$  is invariant, *i.e.*, for each  $s \in \mathbb{R}$  we have  $\Phi_{s*}^L \mu = \mu$ . Let  $f$  be any continuous function on  $\mathcal{E}(E)$ . Then :

$$(12) \quad \left| \int f d\Phi_{s*}^L \mu - \int f d\mu \right| \leq \left| \int f d\Phi_{s*}^L \mu - \int f d\Phi_{s*}^L \mu_{T_n} \right| + \left| \int f d\Phi_{s*}^L \mu_{T_n} - \int f d\mu_{T_n} \right| + \left| \int f d\mu_{T_n} - \int f d\mu \right|.$$

From the weak\* convergence  $\mu_{T_n} \rightarrow \mu$ , we have that

$$\left| \int f d\Phi_{s*}^L \mu_{T_n} - \int f d\mu \right| \rightarrow 0 \quad \text{as } T_n \rightarrow +\infty$$

and

$$\left| \int f d\Phi_{s*}^L \mu - \int f d\Phi_{s*}^L \mu_{T_n} \right| = \left| \int f \circ \Phi_s^L d\mu - \int f \circ \Phi_s^L d\mu_{T_n} \right| \rightarrow 0 \quad \text{as } T_n \rightarrow +\infty.$$

Moreover, using the definition of  $\mu_{T_n}$ :

$$\begin{aligned} \left| \int f d\Phi_{s*}^L \mu_{T_n} - \int f d\mu_{T_n} \right| &= \frac{1}{T_n} \left| \int_s^{T_n+s} f(\gamma(t), \dot{\gamma}(t)) dt - \int_0^{T_n} f(\gamma(t), \dot{\gamma}(t)) dt \right| = \\ &= \frac{1}{T_n} \left| \int_{T_n}^{T_n+s} f(\gamma(t), \dot{\gamma}(t)) dt - \int_0^s f(\gamma(t), \dot{\gamma}(t)) dt \right| \leq \\ &\leq \frac{2s}{T_n} \max_{\mathcal{E}(E)} |f| \rightarrow 0 \quad \text{as } T_n \rightarrow +\infty. \end{aligned}$$

Therefore, it follows from (12) that for any  $f \in \mathcal{E}(E)$ :

$$\left| \int f d\Phi_{s*}^L \mu - \int f d\mu \right| = 0.$$

Consequently,  $\Phi_{s*}^L \mu = \mu$  for any  $s \in \mathbb{R}$ , *i.e.*,  $\mu$  is  $\Phi^L$ -invariant.  $\square$

To each  $\mu \in \mathfrak{M}(L)$ , we may associate its *average action*

$$A_L(\mu) = \int_{TM} L d\mu.$$

Clearly, the measures constructed in Proposition 4.1 have finite action.

**Remark 4.2.** One can show the existence of invariant probability measures with finite action, also in the case of non-autonomous time-periodic Lagrangians. As it was originally done by Mather in [42], one can apply Kryloff and Bogoliouboff's result to a one-point compactification of  $TM$  and consider the extended Lagrangian system that leaves the point at infinity fixed. The main step consists in showing that the measure provided by this construction has no atomic part supported at  $\infty$  (which is a fixed point for the extended system).

**Note:** We shall hereafter assume that  $\mathfrak{M}(L)$  is endowed with the *vague topology*, *i.e.*, the weak\* topology induced by the space  $C_\ell^0$  of continuous functions  $f : TM \rightarrow \mathbb{R}$  having at most linear growth:

$$\sup_{(x,v) \in TM} \frac{|f(x,v)|}{1 + \|v\|} < +\infty.$$

It is not difficult to check that  $\mathfrak{M}(L) \subset (C_\ell^0)^*$ . Moreover it is a classical result that this space, with such a topology, is metrizable. A metric, for instance, can be defined as follows. Let  $\{f_n\}_n$  be a sequence of functions with compact support in

$C_\ell^0$  which is dense in the topology of uniform convergence on compact sets of  $TM$ . Define a metric on  $\mathfrak{M}(L)$  by:

$$d(\mu_1, \mu_2) = \left| \int |v| d\mu_1 - \int |v| d\mu_2 \right| + \sum_n \frac{1}{2^n \|f_n\|_\infty} \left| \int f_n d\mu_1 - \int f_n d\mu_2 \right|.$$

See also Addendum 4.A at the end of this section for more details on this topology.

**Proposition 4.3.**  $A_L : \mathfrak{M}(L) \rightarrow \mathbb{R}$  is lower semicontinuous with the vague topology on  $\mathfrak{M}(L)$ .

**Proof.** Let  $A_{L,K}(\mu) := \int \min\{L, K\} d\mu$ , for  $K \in \mathbb{R}$ . Then,  $A_{L,K}$  is clearly continuous, actually Lipschitz with constant  $K$ . In fact:

$$|A_{L,K}(\mu) - A_{L,K}(\nu)| \leq K d(\mu, \nu).$$

Moreover, since  $A_{L,K} \uparrow A_L$  as  $K \rightarrow \infty$ , it follows that  $A_L$  is lower semicontinuous.  $\square$

**Remark 4.4.** In general, this functional might not be necessarily continuous.

This Proposition has an immediate, but important, consequence.

**Corollary 4.5.** There exists  $\mu \in \mathfrak{M}(L)$ , which minimizes  $A_L$  over  $\mathfrak{M}(L)$ .

We shall call a measure  $\mu \in \mathfrak{M}(L)$ , such that  $A_L(\mu) = \min_{\mathfrak{M}(L)} A_L$ , an *action-minimizing measure* of  $L$  (compare with Remark 3.5 (ii)).

Actually, one can find many other “interesting” measures, besides those found by minimizing  $L$ . We shall see later (see Remark 4.26), that in fact these measures minimizing the action of  $L$ , correspond to special measures with “trivial” homology (or rotation vector). On the other hand, as we have already pointed out in Proposition 3.4 (and the following Warning), in order to get information on a specific KAM torus and characterize it via the invariant probability measures supported on it, one needs necessarily to modify the Lagrangian. Think for instance of an integrable system, for which we have a whole family of KAM tori. It would be unreasonable to expect to obtain ALL of them, just by minimizing a single Lagrangian action. One needs somehow to introduce some sort of “weight”, which, without modifying the dynamics of the system (*i.e.*, the Euler Lagrange flow), allows one to magnify certain motions rather than others. What we found out in Proposition 3.4, was that this can be easily achieved by subtracting to our Lagrangian a linear function  $c \cdot v$ , where  $c$  represented the *cohomology class* of the invariant Lagrangian torus we were interested in, or, in other words, the cohomology class of its graph (that is a closed 1-form) in  $T^*\mathbb{T}^d$ .

A similar idea can be implemented for a general Tonelli Lagrangian. Observe, in fact, that if  $\eta$  is a 1-form on  $M$ , we can interpret it as a function on the tangent space (linear on each fiber)

$$\begin{aligned} \hat{\eta} : TM &\longrightarrow \mathbb{R} \\ (x, v) &\longmapsto \langle \eta(x), v \rangle_x \end{aligned}$$

and consider a new Tonelli Lagrangian  $L_\eta := L - \hat{\eta}$ . The associated Hamiltonian will be given by  $H_\eta(x, p) = H(x, \eta(x) + p)$  (it is sufficient to write down the Fenchel-Legendre transform of  $L$ ).

**Lemma 4.6.** If  $\eta$  is closed, then  $L$  and  $L_\eta$  have the same Euler-Lagrange flow on  $TM$ .

**Proof.** Observe that, since  $\eta$  is closed, the variational equations  $\delta[\int L dt] = 0$  and  $\delta[\int (L - \hat{\eta}) dt] = 0$  have the same extremals for the fixed end-point problem and consequently  $L$  and  $L_\eta$  have the same Euler-Lagrange flows.

One can also check this directly, writing down the Euler-Lagrange equations associated to  $L$  and  $L_\eta$  and observe that, since  $\eta$  is closed (*i.e.*,  $d\eta = 0$ ), then

$$\frac{d}{dt} \left( \frac{\partial}{\partial v} \hat{\eta}(\gamma, \dot{\gamma}) \right) = \frac{d}{dt} (\eta(\gamma)) = \frac{\partial}{\partial x} (\hat{\eta}(\gamma, \dot{\gamma})).$$

Let us show in fact that they are equal component by component. If we denote by  $\eta_j$  and  $\gamma_j$  the  $j$ -th components of  $\eta$  and  $\gamma$ , then the  $i$ th component of the left-hand side of the above equation becomes:

$$\sum_{j=1}^d \frac{\partial \eta_i}{\partial x_j}(\gamma) \cdot \dot{\gamma}_j$$

and on the other side:

$$\sum_{j=1}^d \frac{\partial \eta_j}{\partial x_i}(\gamma) \cdot \dot{\gamma}_j.$$

Since  $\eta$  is closed, then  $\frac{\partial \eta_j}{\partial x_i} = \frac{\partial \eta_i}{\partial x_j}$  for each  $i$  and  $j$ . Therefore, the two expressions are the same. Therefore, the term coming from  $\hat{\eta}$  does not give any contribution to the equations.  $\square$

Although the extremals of these two variational problems are the same, this is not generally true for the orbits “minimizing the action” (we shall give a precise definition of “minimizers” later in section 5). We have already given evidence of this in Proposition 3.4. What one can say is that these action-minimizing objects stay the same when we change the Lagrangian by an exact 1-form.

**Proposition 4.7.** *If  $\mu \in \mathfrak{M}(L)$  and  $\eta = df$  is an exact 1-form, then  $\int \hat{d}f d\mu = 0$ .*

The proof is essentially the same as in Section 3 (Property 2), so we omit it. See also [42, Lemma on page 176]. Let us point out that a measure (not necessarily invariant) that satisfies this condition, is called *closed measure* (see also the Addendum at the end of this section).

Thus, for a fixed  $L$ , the minimizing measures will depend only on the de Rham cohomology class  $c = [\eta] \in H^1(M; \mathbb{R})$ . Therefore, instead of studying the action minimizing properties of a single Lagrangian, one can consider a family of such “modified” Lagrangians, parameterized over  $H^1(M; \mathbb{R})$ .

Hereafter, for any given  $c \in H^1(M; \mathbb{R})$ , we shall denote by  $\eta_c$  a closed 1-form with that cohomology class.

**Definition 4.8.** *Let  $\eta_c$  be a closed 1-form of cohomology class  $c$ . Then, if  $\mu \in \mathfrak{M}(L)$  minimizes  $A_{L_{\eta_c}}$  over  $\mathfrak{M}(L)$ , we shall say that  $\mu$  is a  $c$ -action minimizing measure (or  $c$ -minimal measure, or Mather’s measure with cohomology  $c$ ).*

**Remark 4.9.** Observe that the *cohomology class* of an action-minimizing invariant probability measure, is not something intrinsic in the measure itself nor in the dynamics, but it depends on the specific choice of the Lagrangian  $L$ . Changing the Lagrangian  $L \mapsto L - \eta$  by a closed 1-form  $\eta$ , we shall change all the cohomology classes of its action minimizing measures by  $-\eta \in H^1(M; \mathbb{R})$ . Compare also with Remark 4.14 (ii).

One can consider the function on  $H^1(M; \mathbb{R})$ , which associates to each cohomology class  $c$ , *minus* the value of the corresponding minimal action of the modified

Lagrangian  $L_{\eta_c}$  (the *minus* sign is introduced for a convention that might probably become clearer later on):

$$(13) \quad \begin{aligned} \alpha : H^1(M; \mathbb{R}) &\longrightarrow \mathbb{R} \\ c &\longmapsto - \min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu). \end{aligned}$$

This function  $\alpha$  is well-defined (it does not depend on the choice of the representatives of the cohomology classes) and it is easy to see that it is convex. This is generally known as *Mather's  $\alpha$ -function*. As it happened for KAM tori (see Remark 3.5), we shall see in section 5 that the value  $\alpha(c)$  is related to the energy level containing such  $c$ -action minimizing measures (they coincide). See also [13]. Therefore, if we have an integrable Tonelli Hamiltonian  $H(x, p) = h(p)$  it is easy to deduce that  $\alpha(c) = h(c)$ . For this and several other reasons that we shall see later on, this function is sometime called *effective Hamiltonian*. It is interesting to remark that the value of this  $\alpha$ -function coincides with what is called *Mañé's critical value*, which will be introduced later in sections 5 and 6.

We shall denote by  $\mathfrak{M}_c(L)$  the subset of  $c$ -action minimizing measures:

$$\mathfrak{M}_c := \mathfrak{M}_c(L) = \{\mu \in \mathfrak{M}(L) : A_L(\mu) < +\infty \text{ and } A_{L_{\eta_c}}(\mu) = -\alpha(c)\}.$$

We can now define a first important family of invariant sets (see Remark 3.5 (iii)): the *Mather sets*.

**Definition 4.10.** For a cohomology class  $c \in H^1(M; \mathbb{R})$ , we define the Mather set of cohomology class  $c$  as:

$$(14) \quad \widetilde{\mathcal{M}}_c := \bigcup_{\mu \in \mathfrak{M}_c} \text{supp } \mu \subset TM.$$

The projection on the base manifold  $\mathcal{M}_c = \pi(\widetilde{\mathcal{M}}_c) \subseteq M$  is called *projected Mather set* (with cohomology class  $c$ ).

**Remark 4.11.** (i) This set is clearly non-empty and invariant. Moreover, it is also closed. Observe that usually it is defined as the *closure* of the set on the right-hand side. This is indeed the original definition given by Mather in [42]. However, it is not difficult to check that this set is already closed. In fact, since the space of probability measures on  $TM$  is a separable metric space, one can take a countable dense set  $\{\mu_n\}_{n=1}^\infty$  of Mather's measures and consider the new measure  $\tilde{\mu} = \sum_{n=1}^\infty \frac{1}{2^n} \mu_n$ . This is still an invariant probability measure and it is  $c$ -action minimizing (it follows from the convexity of  $\alpha$ ), hence  $\text{supp } \tilde{\mu} \subseteq \widetilde{\mathcal{M}}_c$ . But, from the definition of  $\tilde{\mu}$  one can clearly deduce that  $\text{supp } \tilde{\mu} = \bigcup_{\mu \in \mathfrak{M}_c} \text{supp } \mu \supseteq \widetilde{\mathcal{M}}_c$ . This implies that  $\text{supp } \tilde{\mu} = \widetilde{\mathcal{M}}_c$ . Therefore, as the support of a single action minimizing measure,  $\widetilde{\mathcal{M}}_c$  is closed (the support of a measure, by definition, is closed). Moreover, this remark points out that there always exists a Mather's measure  $\mu_c$  of full support, *i.e.*,  $\text{supp } \mu_c = \widetilde{\mathcal{M}}_c$ . (ii) As we have pointed out in Remark 3.5, if there is a KAM torus  $\mathcal{T}$  of cohomology class  $c$ , then  $\widetilde{\mathcal{M}}_c = \mathcal{L}^{-1}(\mathcal{T})$ . In particular, the same proof continues to hold if we replace  $\mathcal{T}$  with any invariant Lagrangian graph  $\Lambda$  of cohomology class  $c$ , which supports an invariant measure  $\mu$  of full support (*i.e.*,  $\text{supp } \mu = \Lambda$ ).

In [42] Mather proved the celebrated *graph theorem*:

**Theorem 4.12 (Mather's graph theorem).** Let  $\widetilde{\mathcal{M}}_c$  be defined as in (14). The set  $\widetilde{\mathcal{M}}_c$  is compact, invariant under the Euler-Lagrange flow and  $\pi|_{\widetilde{\mathcal{M}}_c}$  is an injective mapping of  $\widetilde{\mathcal{M}}_c$  into  $M$ , and its inverse  $\pi^{-1} : \mathcal{M}_c \longrightarrow \widetilde{\mathcal{M}}_c$  is Lipschitz.

We shall not prove this theorem here, but it will be deduced from a more general result in Section 5, namely, the graph property of the *Aubry set*. Moreover, similarly to what we have seen for KAM tori (see Remark 3.5), in the autonomous case this set is contained in a well-defined energy level, which can be characterized in terms of the minimal action  $\alpha(c)$ .

**Theorem 4.13 (Carneiro, [13]).** *The Mather set  $\widetilde{\mathcal{M}}_c$  is contained in the energy level  $\{H \circ \mathcal{L}(x, v) = \alpha(c)\}$ .*

Also this result will be deduced from a similar result for the Aubry and Mañé sets. See Proposition 5.23 in Section 5.

Now, we would like to shift our attention to a related problem. As we have seen in section 3, instead of considering different minimizing problems over  $\mathfrak{M}(L)$ , obtained by modifying the Lagrangian  $L$ , one can alternatively try to minimize the Lagrangian  $L$  putting some “constraints”, such as, for instance, fixing the *rotation vector* of the measures. In order to generalize this to Tonelli Lagrangians on compact manifolds, we first need to define what we mean by rotation vector of an invariant measure.

Let  $\mu \in \mathfrak{M}(L)$ . Thanks to the superlinearity of  $L$ , the integral  $\int_{TM} \hat{\eta} d\mu$  is well defined and finite for any closed 1-form  $\eta$  on  $M$ . Moreover, we have proved in Proposition 4.7 that if  $\eta$  is exact, then such an integral is zero, *i.e.*,  $\int_{TM} \hat{\eta} d\mu = 0$ . Therefore, one can define a linear functional:

$$\begin{aligned} H^1(M; \mathbb{R}) &\longrightarrow \mathbb{R} \\ c &\longmapsto \int_{TM} \hat{\eta} d\mu, \end{aligned}$$

where  $\eta$  is any closed 1-form on  $M$  with cohomology class  $c$ . By duality, there exists  $\rho(\mu) \in H_1(M; \mathbb{R})$  such that

$$\int_{TM} \hat{\eta} d\mu = \langle c, \rho(\mu) \rangle \quad \forall c \in H^1(M; \mathbb{R})$$

(the bracket on the right-hand side denotes the canonical pairing between cohomology and homology). We call  $\rho(\mu)$  the *rotation vector* of  $\mu$  (compare with the definition given in Section 3 Property (1)). This rotation vector is the same as the Schwartzman's asymptotic cycle of  $\mu$  (see [52] for more details).

**Remark 4.14.** (i) It is possible to provide a more “geometrical” interpretation of this. Suppose for the moment that  $\mu$  is ergodic. Then, it is known that a generic orbit  $\gamma(t) := \pi \Phi_t^L(x, v)$ , where  $\pi : TM \longrightarrow M$  denotes the canonical projection, will return infinitely many often close (as close as we like) to its initial point  $\gamma(0) = x$ . We can therefore consider a sequence of times  $T_n \rightarrow +\infty$  such that  $d(\gamma(T_n), x) \rightarrow 0$  as  $n \rightarrow +\infty$ , and consider the closed loops  $\sigma_n$  obtained by “closing”  $\gamma|_{[0, T_n]}$  with the shortest geodesic connecting  $\gamma(T_n)$  to  $x$ . Denoting by  $[\sigma_n]$  the homology class of this loop, one can verify [52] that  $\lim_{n \rightarrow \infty} \frac{[\sigma_n]}{T_n} = \rho(\mu)$ , independently of the chosen sequence  $\{T_n\}_n$ . In other words, in the case of ergodic measures, the rotation vector tells us how on average a generic orbit winds around  $TM$ . If  $\mu$  is not ergodic,  $\rho(\mu)$  loses this neat geometric meaning, yet it may be interpreted as the average of the rotation vectors of its different ergodic components.

(ii) It is clear from the discussion above that the rotation vector of an invariant measure depends only on the dynamics of the system (*i.e.*, the Euler-Lagrange flow) and not on the chosen Lagrangian. Therefore, it does not change when we modify our Lagrangian adding a closed one form.

A natural question is whether or not for a given Tonelli Lagrangian  $L$ , there exist invariant probability measures for any given rotation vector. The answer turns out to be affirmative.

In fact, using that the action functional  $A_L : \mathfrak{M}(L) \rightarrow \mathbb{R}$  is lower semicontinuous, one can prove the following [42]:

**Proposition 4.15.** (i) *The map  $\rho : \mathfrak{M}(L) \rightarrow H_1(M; \mathbb{R})$  is continuous.*  
(ii) *The map  $\rho : \mathfrak{M}(L) \rightarrow H_1(M; \mathbb{R})$  is affine, i.e., for any  $\mu, \nu \in \mathfrak{M}(L)$  and  $a, b \geq 0$  with  $a + b = 1$ ,  $\rho(a\mu + b\nu) = a\rho(\mu) + b\rho(\nu)$ .*  
(iii) *For every  $h \in H_1(M; \mathbb{R})$  there exists  $\mu \in \mathfrak{M}(L)$  with  $A_L(\mu) < \infty$  and  $\rho(\mu) = h$ . In other words, the map  $\rho$  is surjective.*

**Proof.** (i) Let us fix a basis  $h_1, \dots, h_b$  in  $H_1(M; \mathbb{R})$ , where  $b$  is the first Betti number of  $M$ , i.e.,  $b = \dim H_1(M; \mathbb{R})$ . If  $\mu_n \rightarrow \mu$  and  $\eta_c$  is any closed one-form on  $M$  of cohomology class  $c$ , then

$$\langle c, \rho(\mu_n) \rangle = \int \eta_c \cdot v d\mu_n \xrightarrow{n \rightarrow +\infty} \int \eta_c \cdot v d\mu = \langle c, \rho(\mu) \rangle,$$

that is equivalent to say that:

$$\langle c, \rho(\mu_n) - \rho(\mu) \rangle \xrightarrow{n \rightarrow +\infty} 0 \quad \forall c \in \mathbb{R}^b \simeq H^1(M; \mathbb{R}).$$

Therefore,  $\rho(\mu_n) - \rho(\mu) \rightarrow 0$  as  $n \rightarrow +\infty$ .

(ii) The fact that the map  $\rho$  is affine, is a trivial consequence of the definition of rotation vector.

(iii) Let  $h \in H_1(M; \mathbb{Z})$  be an integer homology class and choose  $\gamma_h : [0, 1] \rightarrow M$  a closed loop with homology  $h$ . Let us consider the loop space:

$$\Omega_h := \{\xi : [0, 1] \rightarrow M, \text{ abs. cont. loop with the same free homotopy class as } \gamma_h\}.$$

One can prove [35] that the Lagrangian action functional  $A_L$  has a minimum on this space, which is a periodic orbit of the Euler-Lagrange flow, with period 1 (it is essentially Tonelli's theorem). Let us consider this periodic orbit  $\gamma : [0, 1] \rightarrow M$  and define an invariant probability measure  $\mu_\gamma$  evenly distributed along this periodic orbit:

$$\int f d\mu_\gamma := \int_0^1 f(\gamma(t), \dot{\gamma}(t)) dt \quad \forall f \in C^0(TM).$$

It is easy to verify that such a measure is invariant, i.e.,  $\Phi_t^{L*} \mu_\gamma = \mu_\gamma$  and, using the definition of rotation vector and Remark 4.14, that  $\rho(\mu_\gamma) = h$ . Since the map  $\rho$  is affine and  $\mathfrak{M}(L)$  is convex, it follows that  $\rho(\mathfrak{M}(L))$  is convex. It follows from what discussed above, that this set must contain the convex hull of  $H_1(M; \mathbb{Z})$  and therefore  $\rho(\mathfrak{M}(L)) = H_1(M; \mathbb{R})$ .  $\square$

As already pointed out in Section 3, amongst all probability measures with a prescribed rotation vector, a peculiar role - from a dynamical systems point of view - will be played by those minimizing the average action. Following Mather, let us consider the minimal value of the average action  $A_L$  over the probability measures with rotation vector  $h$ . Observe that this minimum is actually achieved because of the lower semicontinuity of  $A_L$  and the compactness of  $\rho^{-1}(h)$  ( $\rho$  is continuous and  $L$  superlinear). Let us define

$$(15) \quad \begin{aligned} \beta : H_1(M; \mathbb{R}) &\longrightarrow \mathbb{R} \\ h &\longmapsto \min_{\mu \in \mathfrak{M}(L) : \rho(\mu) = h} A_L(\mu). \end{aligned}$$

This function  $\beta$  is what is generally known as *Mather's  $\beta$ -function* and it is immediate to check that it is convex. As we have noticed in Remark 3.7, if there is a

KAM torus of cohomology class  $c$  and rotation vector  $\rho$ , then  $\beta(\rho) = -E_c + c \cdot \rho$ . Therefore, if we have an integrable Tonelli Hamiltonian  $H(x, p) = h(p)$  and the associated Lagrangian  $L(x, v) = \ell(v)$ , it is easy to deduce that  $\beta(h) = \ell(h)$ . For this and several other reasons that we shall see later on, this function is sometime called *effective Lagrangian*. Moreover, this function is also related to the notion of *stable norm* for a metric  $d$  (see for instance [36]).

We can now define what we mean by action minimizing measure with a given rotation vector (compare with Remark 3.7(ii)).

**Definition 4.16.** A measure  $\mu \in \mathfrak{M}(L)$  realizing the minimum in (15), i.e., such that  $A_L(\mu) = \beta(\rho(\mu))$ , is called an *action minimizing* (or *minimal* or *Mather's*) *measure with rotation vector*  $\rho(\mu)$ .

**Remark 4.17.** We shall see in Section 5 that, differently from what happens with invariant probability measures, it will not be always possible to find *action-minimizing orbits* for any given rotation vector (not even define a rotation vector for each action minimizing orbit). This is one of the main difference with the twist map case. In higher dimensions, in fact, an example due to Hedlund [28] provides the existence of a Riemannian metric on a three-dimensional torus, for which minimal geodesics exist only in three directions.

We shall denote by  $\mathfrak{M}^h(L)$  the subset of action minimizing measures with rotation vector  $h$ :

$$\mathfrak{M}^h := \mathfrak{M}^h(L) = \{\mu \in \mathfrak{M}(L) : A_L(\mu) < +\infty, \rho(\mu) = h \text{ and } A_L(\mu) = \beta(h)\}.$$

This allows us to define another important family of invariant sets (see also Remark 3.7 (iii)).

**Definition 4.18.** For a homology class (or rotation vector)  $h \in H_1(M; \mathbb{R})$ , we define the *Mather set corresponding to a rotation vector*  $h$  as

$$(16) \quad \widetilde{\mathcal{M}}^h := \bigcup_{\mu \in \mathfrak{M}^h} \text{supp } \mu \subset \text{TM},$$

and the *projected one* as  $\mathcal{M}^h = \pi(\widetilde{\mathcal{M}}^h) \subseteq M$ .

**Remark 4.19.** (i) Similarly to what we have pointed out in Remark 4.11, this set is also non-empty and invariant. Moreover, it is also closed. Also in this case, it is not necessary to consider the closure of this set - as it is usually done in the literature - since it is already closed. Just take a countable dense set  $\{\mu_n\}_{n=1}^\infty$  of Mather's measures with rotation vector  $h$  and consider the new measure  $\tilde{\mu} = \sum_{n=1}^\infty \frac{1}{2^n} \mu_n$ . This is still an invariant probability measure and its rotation vector is  $h$ . Moreover, it follows from the convexity of  $\beta$  that this measure is action-minimizing among all measures with rotation vector  $h$ . Hence,  $\text{supp } \tilde{\mu} \subseteq \widetilde{\mathcal{M}}^h$ . Clearly  $\text{supp } \tilde{\mu} = \overline{\bigcup_{\mu \in \mathfrak{M}^h} \text{supp } \mu} \supseteq \widetilde{\mathcal{M}}^h$ . This implies that  $\text{supp } \tilde{\mu} = \widetilde{\mathcal{M}}^h$ . Therefore,  $\widetilde{\mathcal{M}}^h$  is closed. Moreover, this shows that there always exists a Mather's measure  $\mu^h$  of full support, i.e.,  $\text{supp } \mu^h = \widetilde{\mathcal{M}}^h$ .

(ii) As we have pointed out in Remark 3.7, if there is a KAM torus  $\mathcal{T}$  of rotation vector  $h$ , then  $\widetilde{\mathcal{M}}^h = \mathcal{L}^{-1}(\mathcal{T})$ . In particular, the same proof continues to hold if we replace  $\mathcal{T}$  with any invariant Lagrangian graph  $\Lambda$ , which supports an invariant measure  $\mu$  of rotation vector  $\rho$  and of full support (i.e.,  $\text{supp } \mu = \Lambda$ ).

Also for the Mather set corresponding to a rotation vector, we can prove a result similar to Theorem 4.12.

**Theorem 4.20.** *Let  $\widetilde{\mathcal{M}}^h$  be defined as in (16).  $\widetilde{\mathcal{M}}^h$  is compact, invariant under the Euler-Lagrange flow and  $\pi|_{\widetilde{\mathcal{M}}^h}$  is an injective mapping of  $\widetilde{\mathcal{M}}^h$  into  $M$  and its inverse  $\pi^{-1} : \mathcal{M}^h \rightarrow \widetilde{\mathcal{M}}^h$  is Lipschitz.*

**Remark 4.21.** Although the graph property for  $\widetilde{\mathcal{M}}^h$  is not proved in [42], it is easy to deduce it from Theorem 4.12, using the fact that  $\widetilde{\mathcal{M}}^h$  can be seen as the support of a single action-minimizing measure (Remark 4.19) and that this set is included in some  $\widetilde{\mathcal{M}}_c$ , for some suitable  $c \in H^1(M; \mathbb{R})$  (Proposition 4.25).

The above discussion leads to two equivalent formulations for the minimality of a measure  $\mu$ :

- there exists a homology class  $h \in H_1(M; \mathbb{R})$ , namely its rotation vector  $\rho(\mu)$ , such that  $\mu$  minimizes  $A_L$  amongst all measures in  $\mathfrak{M}(L)$  with rotation vector  $h$ ; i.e.,  $A_L(\mu) = \beta(h)$ .
- There exists a cohomology class  $c \in H^1(M; \mathbb{R})$ , such that  $\mu$  minimizes  $A_{L_{\eta_c}}$  amongst all probability measures in  $\mathfrak{M}(L)$ ; i.e.,  $A_{L_{\eta_c}}(\mu) = -\alpha(c)$ .

What is the relation between two these different approaches? Are they equivalent, i.e.,  $\bigcup_{h \in H_1(M; \mathbb{R})} \mathfrak{M}^h = \bigcup_{c \in H^1(M; \mathbb{R})} \mathfrak{M}_c$ ?

In order to comprehend the relation between these two families of action-minimizing measures, we need to understand better the properties of these functions

$$\alpha : H^1(M; \mathbb{R}) \rightarrow \mathbb{R} \quad \text{and} \quad \beta : H_1(M; \mathbb{R}) \rightarrow \mathbb{R}$$

that we have introduced above. Let us start with the following trivial remark.

**Remark 4.22.** As we have previously pointed out, if we have an integrable Tonelli Hamiltonian  $H(x, p) = h(p)$  and the associated Lagrangian  $L(x, v) = \ell(v)$ , then  $\alpha(c) = h(c)$  and  $\beta(h) = \ell(h)$ . In this case, the cotangent space  $T^*\mathbb{T}^d$  is foliated by invariant tori  $\mathcal{T}_c^* := \mathbb{T}^d \times \{c\}$  and the tangent space  $T\mathbb{T}^d$  by invariant tori  $\widetilde{\mathcal{T}}^h := \mathbb{T}^d \times \{h\}$ . In particular, we proved that

$$\widetilde{\mathcal{M}}_c = \mathcal{L}^{-1}(\mathcal{T}_c) = \widetilde{\mathcal{T}}^h = \widetilde{\mathcal{M}}^h,$$

where  $h$  and  $c$  are such that  $h = \nabla h(c) = \nabla \alpha(c)$  and  $c = \nabla \ell(h) = \nabla \beta(h)$ .

We would like to prove that such a relation that links Mather sets of a certain cohomology class to Mather sets with a given rotation vector, goes beyond the specificity of this situation. Of course, one main difficulty is that in general the *effective Hamiltonian*  $\alpha$  and the *effective Lagrangian*  $\beta$ , although being convex and superlinear (see Proposition 4.23), are not necessarily differentiable.

Before stating and proving the main relation between these two functions, let us recall some definitions and results from classical convex analysis (see [51]). Given a convex function  $\varphi : V \rightarrow \mathbb{R} \cup \{+\infty\}$  on a finite dimensional vector space  $V$ , one can consider a *dual* (or *conjugate*) function defined on the dual space  $V^*$ , via the so-called *Fenchel transform*:  $\varphi^*(p) := \sup_{v \in V} (p \cdot v - \varphi(v))$ .

**Proposition 4.23.**  *$\alpha$  and  $\beta$  are convex conjugate, i.e.,  $\alpha^* = \beta$  and  $\beta^* = \alpha$ . In particular, it follows that  $\alpha$  and  $\beta$  have superlinear growth.*



**Proof.** First of all, recall that  $(H_1(M; \mathbb{R}))^* \simeq H^1(M; \mathbb{R})$  and  $(H^1(M; \mathbb{R}))^* \simeq H_1(M; \mathbb{R})$ . Let us compute  $\beta^*$ :

$$\begin{aligned}
 \beta^*(c) &= \max_{h \in H_1(M; \mathbb{R})} (\langle c, h \rangle - \beta(h)) = \\
 &= - \min_{h \in H_1(M; \mathbb{R})} (\beta(h) - \langle c, h \rangle) = \\
 &= - \min_{h \in H_1(M; \mathbb{R})} \left( \min_{\mu \in \mathfrak{M}^h(L)} A_L(\mu) - \langle c, h \rangle \right) = \\
 &= - \min_{h \in H_1(M; \mathbb{R})} \min_{\mu \in \mathfrak{M}^h(L)} (A_L(\mu) - \langle c, \rho(\mu) \rangle) = \\
 &= - \min_{h \in H_1(M; \mathbb{R})} \min_{\mu \in \mathfrak{M}^h(L)} A_{L_{\eta_c}}(\mu) = \\
 &= - \min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu) = \\
 &= -\alpha(c).
 \end{aligned}$$

Similarly, one can check that  $\alpha^* = \beta$ :

$$\begin{aligned}
 \alpha^*(h) &= \max_{c \in H^1(M; \mathbb{R})} (\langle c, h \rangle - \alpha(c)) = \\
 &= \max_{c \in H^1(M; \mathbb{R})} \left( \langle c, h \rangle + \min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu) \right) = \\
 &= \max_{c \in H^1(M; \mathbb{R})} \min_{\mu \in \mathfrak{M}(L)} (A_L(\mu) + \langle c, h - \rho(\mu) \rangle) = \\
 &= \min_{\mu \in \mathfrak{M}(L)} \max_{c \in H^1(M; \mathbb{R})} (A_L(\mu) + \langle c, h - \rho(\mu) \rangle) = \\
 &= \min_{\rho \in H_1(M; \mathbb{R})} \min_{\mu \in \mathfrak{M}^\rho(L)} \max_{c \in H^1(M; \mathbb{R})} (A_L(\mu) + \langle c, h - \rho \rangle),
 \end{aligned}$$

where in the second last line we could exchange the order of the max and the min, using a general result by Rockafellar, that requires concavity in one variable, convexity in the other one, and some compactness assumption (see [51, Section 36]). Observe now that if  $h \neq \rho$ , then  $\max_{c \in H^1(M; \mathbb{R})} (A_L(\mu) + \langle c, h - \rho \rangle) = +\infty$ . Therefore:

$$\begin{aligned}
 \alpha^*(h) &= \dots = \min_{\rho \in H_1(M; \mathbb{R})} \min_{\mu \in \mathfrak{M}^\rho(L)} \max_{c \in H^1(M; \mathbb{R})} (A_L(\mu) + \langle c, h - \rho \rangle) = \\
 &= \min_{\mu \in \mathfrak{M}^h(L)} A_L(\mu) = \beta(h).
 \end{aligned}$$

The second statement of this proposition follows from a general property of convex conjugation. Let  $\varphi : V \rightarrow \mathbb{R}$  be a convex function on a finite dimensional vector space  $V$  and let  $\varphi^* : V^* \rightarrow \mathbb{R} \cup \{+\infty\}$  be its convex conjugate. Then (see [51]):  $\varphi^*$  is finite everywhere if and only if  $\varphi$  has superlinear growth, i.e.,  $\frac{\varphi(x)}{\|x\|} \rightarrow +\infty$  as  $\|x\| \rightarrow +\infty$ . □

Next proposition will allow us to clarify the relation (and duality) between the two minimizing procedures above. To state it, recall that, like any convex function on a finite-dimensional space,  $\beta$  admits a subderivative at each point  $h \in H_1(M; \mathbb{R})$ , i.e., we can find  $c \in H^1(M; \mathbb{R})$  such that

$$(17) \quad \forall h' \in H_1(M; \mathbb{R}), \quad \beta(h') - \beta(h) \geq \langle c, h' - h \rangle.$$

As it is usually done, we shall denote by  $\partial\beta(h)$  the set of  $c \in H^1(M; \mathbb{R})$  that are subderivatives of  $\beta$  at  $h$ , i.e., the set of  $c$  which satisfy the inequality above. Similarly, we shall denote by  $\partial\alpha(c)$  the set of subderivatives of  $\alpha$  at  $c$ . Fenchel's duality implies an easier characterization of subdifferentials.

**Proposition 4.24.**  *$c \in \partial\beta(h)$  if and only if  $\langle c, h \rangle = \alpha(c) + \beta(h)$ . Similarly  $h \in \partial\alpha(c)$  if and only if  $\langle c, h \rangle = \alpha(c) + \beta(h)$ . In particular,  $c \in \partial\beta(h)$  if and only if  $h \in \partial\alpha(c)$ .*

**Proof.** We shall prove only the first statement. The second one is analogous and the third one a trivial consequence.

[ $\Leftarrow$ ] Using that  $\beta(h') \geq \langle c, h' \rangle - \alpha(c)$  for each  $h' \in H_1(M; \mathbb{R})$ , it follows that for each  $h' \in H_1(M; \mathbb{R})$ :

$$\begin{aligned} \beta(h') - \beta(h) &\geq \langle c, h' \rangle - \alpha(c) - \beta(h) = \langle c, h' \rangle - (\alpha(c) + \beta(h)) = \\ &= \langle c, h' \rangle - \langle c, h \rangle = \langle c, h' - h \rangle. \end{aligned}$$

[ $\Rightarrow$ ] It follows from Fenchel-Legendre inequality that  $\alpha(c) + \beta(h') \geq \langle c, h' \rangle$  for each  $h' \in H_1(M; \mathbb{R})$ . Therefore we only need to prove the reverse inequality. In fact, using (17) one can deduce that for each  $h' \in H_1(M; \mathbb{R})$ :

$$\begin{aligned} \alpha(c) + \beta(h) &\leq \alpha(c) + \beta(h') - \langle c, h' - h \rangle = \alpha(c) + \beta(h') - \langle c, h' \rangle + \langle c, h \rangle = \\ &= \left( \alpha(c) + \beta(h') - \langle c, h' \rangle \right) + \langle c, h \rangle. \end{aligned}$$

Therefore, taking the minimum over  $h'$  on the left-hand side we obtain:

$$\alpha(c) + \beta(h) \leq \min_{h' \in H_1(M; \mathbb{R})} \left( \alpha(c) + \beta(h') - \langle c, h' \rangle \right) + \langle c, h \rangle = \langle c, h \rangle,$$

where in the last equality we used that  $\min_{h' \in H_1(M; \mathbb{R})} (\alpha(c) + \beta(h') - \langle c, h' \rangle) = 0$ , as it follows easily from Fenchel-Legendre duality.  $\square$

We can now prove that what observed in Remark 4.22 continues to hold in the general case

**Proposition 4.25.** *Let  $\mu \in \mathfrak{M}(L)$  be an invariant probability measure. Then:*

- (i)  *$A_L(\mu) = \beta(\rho(\mu))$  if and only if there exists  $c \in H^1(M; \mathbb{R})$  such that  $\mu$  minimizes  $A_{L_{\eta_c}}$  (i.e.,  $A_{L_{\eta_c}}(\mu) = -\alpha(c)$ ).*
- (ii) *If  $\mu$  satisfies  $A_L(\mu) = \beta(\rho(\mu))$  and  $c \in H^1(M; \mathbb{R})$ , then  $\mu$  minimizes  $A_{L_{\eta_c}}$  if and only if  $c \in \partial\beta(\rho(\mu))$  (or equivalently  $\langle c, h \rangle = \alpha(c) + \beta(\rho(\mu))$ ).*

**Proof.** We shall prove both statements at the same time. Assume  $A_L(\mu_0) = \beta(\rho(\mu_0))$ . Let  $c \in \partial\beta(\rho(\mu_0))$ , by Fenchel's duality this is equivalent to

$$\begin{aligned} \alpha(c) &= \langle c, \rho(\mu_0) \rangle - \beta(\rho(\mu_0)) \\ &= \langle c, \rho(\mu_0) \rangle - A_L(\mu_0) \\ &= -A_{L_{\eta_c}}(\mu_0). \end{aligned}$$

Therefore  $A_{L_{\eta_c}}(\mu_0) = \min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu)$ .

Assume conversely that  $A_{L_{\eta_c}}(\mu_0) = \min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu)$ , for some given cohomology class  $c$ . Then, it follows that

$$\alpha(c) = -A_{L_{\eta_c}}(\mu_0),$$

which can be written as

$$\langle c, \rho(\mu_0) \rangle = \alpha(c) + A_L(\mu_0).$$

It now suffices to use the Fenchel inequality  $\langle c, \rho(\mu_0) \rangle \leq \alpha(c) + \beta(\rho(\mu_0))$ , and the inequality  $\beta(\rho(\mu_0)) \leq A_L(\mu_0)$ , given by the definition of  $\beta$ , to obtain the equality

$$\langle c, \rho(\mu_0) \rangle = \alpha(c) + \beta(\rho(\mu_0)).$$

In particular, we have  $A_L(\mu_0) = \beta(\rho(\mu_0))$ .  $\square$

**Remark 4.26.** (i) It follows from the above proposition, that both minimizing procedures lead to the same sets of invariant probability measures:

$$\bigcup_{h \in H_1(M; \mathbb{R})} \mathfrak{M}^h = \bigcup_{c \in H^1(M; \mathbb{R})} \mathfrak{M}_c.$$

In other words, minimizing over the set of invariant measures with a fixed rotation vector or minimizing - globally - the modified Lagrangian (corresponding to a certain cohomology class) are dual problems, as the ones that often appears in linear programming and optimization.

(ii) In particular, we have the following inclusions between Mather sets:

$$c \in \partial\beta(h) \iff h \in \partial\alpha(c) \iff \widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_c.$$

Moreover, for any  $c \in H^1(M; \mathbb{R})$ :

$$\widetilde{\mathcal{M}}_c = \bigcup_{h \in \partial\alpha(c)} \widetilde{\mathcal{M}}^h.$$

(iii) The minimum of the  $\alpha$  function is sometime called *Mañé's strict critical value*. Observe that if  $\alpha(c_0) = \min \alpha(c)$ , then  $0 \in \partial\alpha(c_0)$  and  $\beta(0) = -\alpha(c_0)$ . Therefore, the measures with zero homology are contained in the least possible energy level containing Mather sets:  $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_{c_0}$ . This inclusion might be strict, unless  $\alpha$  is differentiable at  $c_0$ ; in fact, there may be other action minimizing measures with non-zero rotation vectors corresponding to the other subderivatives of  $\alpha$  at  $c_0$ .

(iv) Note that measures of trivial homology are not necessarily supported on orbits with trivial homology or fixed points. For instance, one can consider the following example (see also [16, Section 5.2]). Let  $M = \mathbb{T}^2$  equipped with the flat metric and consider a vector field  $X$  with norm 1 and such that its orbits form a Reeb foliation, *i.e.*,  $X$  has two closed orbits  $\gamma_1$  and  $\gamma_2$  in opposite homology classes and any other orbit asymptotically approaches  $\gamma_1$  in forward time and  $\gamma_2$  in backward time. As we have described in section 2, we can embed this vector field into the Euler-Lagrange vector field given by the Tonelli Lagrangian  $L_X(x, v) = \frac{1}{2} \|v - X(x)\|^2$ . Let us now consider the probability measure  $\mu_{\gamma_1}$  and  $\mu_{\gamma_2}$ , uniformly distributed respectively on  $\gamma_1$  and  $\gamma_2$ . Since these two curves have opposite homologies, then  $\rho(\mu_{\gamma_1}) = -\rho(\mu_{\gamma_2}) =: h_0 \neq 0$ . Moreover, it is easy to see that  $A_{L_X}(\mu_{\gamma_1}) = A_{L_X}(\mu_{\gamma_2}) = 0$ , since the Lagrangian vanishes on  $\text{Graph}(X)$ . Using the fact that  $L_X \geq 0$  (in particular it is strictly positive outside of  $\text{Graph}(X)$ ) and that there are no other invariant ergodic probability measures contained in  $\text{Graph}(X)$ , we can conclude that  $\mathcal{M}_0 = \gamma_1 \cup \gamma_2$  and  $\alpha(0) = 0$ . Moreover,  $\mu_0 := \frac{1}{2}\mu_{\gamma_1} + \frac{1}{2}\mu_{\gamma_2}$  has zero homology and its support is contained in  $\widetilde{\mathcal{M}}_0$ . Therefore (see Proposition 4.25 (i)),  $\mu_0$  is action minimizing with rotation vector 0 and  $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_0$ ; in particular,  $\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0$ . This also implies that  $\beta(0) = 0$  and  $\alpha(0) = \min \alpha(c) = 0$ .

Observe that  $\alpha$  is not differentiable at 0. In fact, reasoning as we have done before for the zero homology class, it is easy to see that for all  $t \in [-1, 1]$   $\widetilde{\mathcal{M}}^{th_0} = \widetilde{\mathcal{M}}_0$ . It is sufficient to consider the convex combination  $\mu_\lambda = \lambda\mu_{\gamma_1} + (1 - \lambda)\mu_{\gamma_2}$  for any  $\lambda \in [0, 1]$ . Therefore,  $\partial\alpha(0) = \{th_0, t \in [-1, 1]\}$  and  $\beta(th_0) = 0$  for all  $t \in [-1, 1]$ .

As we have just seen in item (iv) of Remark 4.26, it may happen that the Mather sets corresponding to different homology (resp. cohomology) classes coincide or are included one into the other. This is something that, for instance, cannot happen in the integrable case: in this situation, in fact, these sets form a foliation and are disjoint. The problem in the above mentioned example, seems to be related to a lack of *strict convexity* of  $\beta$  and  $\alpha$ . See also the discussion on the simple pendulum

in Addendum 4.B: in this case the Mather sets, corresponding to a non-trivial interval of cohomology classes about 0, coincide.

In the light of this, let us try to understand better what happens when  $\alpha$  and  $\beta$  are not strictly convex, *i.e.*, when we are in the presence of “*flat*” pieces.

Let us first fix some notation. If  $V$  is a real vector space and  $v_0, v_1 \in V$ , we shall denote by  $\sigma(v_0, v_1)$  the segment joining  $v_0$  to  $v_1$ , that is  $\sigma(v_0, v_1) := \{tv_0 + (1-t)v_1 : t \in [0, 1]\}$ . We shall say that a function  $f : V \rightarrow \mathbb{R}$  is *affine* on  $\sigma(v_0, v_1)$ , if there exists  $v^* \in V^*$  (the dual of  $V$ ), such that  $f(v) = f(v_0) + \langle v^*, v - v_0 \rangle$  for each  $v \in \sigma(v_0, v_1)$ . Moreover, we shall denote by  $\text{Int}(\sigma(v_0, v_1))$  the *interior* of  $\sigma(v_0, v_1)$ , *i.e.*,  $\text{Int}(\sigma(v_0, v_1)) := \{tv_0 + (1-t)v_1 : t \in (0, 1)\}$ .

**Proposition 4.27.** (i) Let  $h_0, h_1 \in H_1(M; \mathbb{R})$ .  $\beta$  is affine on  $\sigma(h_0, h_1)$  if and only if for any  $h \in \text{Int}(\sigma(h_0, h_1))$  we have  $\widetilde{\mathcal{M}}^h \supseteq \widetilde{\mathcal{M}}^{h_0} \cup \widetilde{\mathcal{M}}^{h_1}$ .

(i) Let  $c_0, c_1 \in H^1(M; \mathbb{R})$ .  $\alpha$  is constant on  $\sigma(c_0, c_1)$  if and only if for any  $c \in \text{Int}(\sigma(c_0, c_1))$  we have  $\widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{M}}_{c_0} \cap \widetilde{\mathcal{M}}_{c_1}$ .

**Remark 4.28.** The inclusions in Proposition 4.26 may not be true at the end points of  $\sigma$ . For instance, Remark 4.26 (iv) provides an example in which the inclusion in Proposition 4.27 (i) is not true at the end-points of  $\sigma(-h_0, h_0)$ .

**Proof.** (i) [ $\implies$ ] Assume that  $\beta$  is affine on  $\sigma(h_0, h_1)$ , *i.e.*, there exists  $c \in H^1(M; \mathbb{R})$  such that  $\beta(h) = \beta(h_0) + \langle c, h - h_0 \rangle$  for all  $h \in \sigma(h_0, h_1)$ .

For  $i = 0, 1$ , let  $\mu_i$  be an action minimizing measure with rotation vector  $h_i$ . If  $t \in (0, 1)$ , let us consider  $\mu_t := t\mu_1 + (1-t)\mu_0$ . Clearly,  $\rho(\mu_t) = th_1 + (1-t)h_0 \in \text{Int}(\sigma(h_0, h_1))$ , therefore  $\beta(\rho(\mu_t)) = \beta(h_0) + t\langle c, h_1 - h_0 \rangle$ . Then:

$$\begin{aligned} \int_{\text{TM}} L(x, v) d\mu_t &= \int_{\text{TM}} L(x, v) d(t\mu_1 + (1-t)\mu_0) = \\ &= t \int_{\text{TM}} L(x, v) d\mu_1 + (1-t) \int_{\text{TM}} L(x, v) d\mu_0 = \\ &= t\beta(h_1) + (1-t)\beta(h_0) = \\ &= t(\beta(h_0) + \langle c, h_1 - h_0 \rangle) + (1-t)\beta(h_0) = \\ &= \beta(h_0) + t\langle c, h_1 - h_0 \rangle = \beta(\rho(\mu_t)). \end{aligned}$$

Therefore  $\mu_t$  is action minimizing with rotation vector  $th_1 + (1-t)h_0$ . Since this is true for any  $\mu_i$  with rotation vector  $h_i$  ( $i = 0, 1$ ), it follows that:

$$\widetilde{\mathcal{M}}^{h_0} \cup \widetilde{\mathcal{M}}^{h_1} \subseteq \widetilde{\mathcal{M}}^{th_1 + (1-t)h_0}.$$

[ $\impliedby$ ] Since  $\widetilde{\mathcal{M}}^h \supseteq \widetilde{\mathcal{M}}^{h_0} \cup \widetilde{\mathcal{M}}^{h_1}$  for all  $h \in \text{Int}(\sigma(h_0, h_1))$ , using the observation in Remark 4.26 (ii) and Proposition 4.25, we obtain that  $\widetilde{\mathcal{M}}^{h_0}$  and  $\widetilde{\mathcal{M}}^{h_1}$  must be contained in  $\widetilde{\mathcal{M}}_{c_0}$  for some  $c_0 \in H^1(M; \mathbb{R})$ . In particular,  $c_0 \in \partial\beta(h_0) \cap \partial\beta(h_1)$ . Let us show that  $c_0 \in \partial\beta(h)$  for all  $h \in \text{Int}(\sigma(h_0, h_1))$ . In fact, let  $h = th_1 + (1-t)h_0$  for some  $t \in (0, 1)$ . Then using the convexity of  $\beta$ :

$$\begin{aligned} \alpha(c_0) + \beta(h) &= \alpha(c_0) + \beta(th_1 + (1-t)h_0) \leq \\ &\leq t(\alpha(c_0) + \beta(h_1)) + (1-t)(\alpha(c_0) + \beta(h_0)) = \\ &\leq t\langle c_0, h_1 \rangle + (1-t)\langle c_0, h_0 \rangle = \\ &= \langle c_0, th_1 + (1-t)h_0 \rangle = \langle c_0, h \rangle. \end{aligned}$$

On the hand, the reverse inequality is always true (Fenchel-Legendre inequality).

Now using this fact, it follows that if  $h \in \sigma(h_0, h_1)$  then:

$$\alpha(c_0) + \beta(h) = \langle c_0, h \rangle \quad \text{and} \quad \alpha(c_0) + \beta(h_0) = \langle c_0, h_0 \rangle$$

therefore subtracting the second equality from the first one, we obtain what we wanted:

$$\beta(h) - \beta(h_0) = \langle c_0, h - h_0 \rangle \quad \forall h \in \sigma(h_0, h_1).$$

(ii) [ $\Leftarrow$ ] It is a trivial consequence of Theorem 4.13.

[ $\Rightarrow$ ] Let  $t \in (0, 1)$  and consider  $c_t := tc_1 + (1-t)c_0$ . We want to show that for each  $h \in \partial\alpha(c_t)$  we have that  $\langle c_1 - c_0, h \rangle = 0$ . First observe that since  $c_t \in \text{Int}(\sigma(c_0, c_1))$ , then there exists  $\delta > 0$  such that  $c_t + s(c_1 - c_0) \in \sigma(c_0, c_1)$  for all  $s \in (-\delta, \delta)$ . Then, using the fact that  $h \in \partial\alpha(c_t)$  and that  $\alpha$  is constant in  $\sigma(c_0, c_1)$ , we obtain:

$$0 = \alpha(c_t + s(c_1 - c_0)) - \alpha(c_t) \geq s\langle c_1 - c_0, h \rangle \quad \forall s \in (-\delta, \delta).$$

But this can be true only if  $\langle c_1 - c_0, h \rangle = 0$ .

Now let us prove that, for  $i = 0, 1$ ,  $\partial\alpha(c_i) \supseteq \partial\alpha(c_t)$  for all  $t \in (0, 1)$ . In fact, if  $h \in \partial\alpha(c_t)$ , then using that  $\langle c_1, h \rangle = \langle c_0, h \rangle$  and that  $\alpha$  is constant on  $\sigma(c_0, c_1)$ , we get:

$$\begin{aligned} \alpha(c_i) + \beta(h) &= \alpha(c_t) + \beta(h) = \langle c_t, h \rangle = \\ &= \langle tc_1 + (1-t)c_0, h \rangle = \\ &= t\langle c_1, h \rangle + (1-t)\langle c_0, h \rangle = \\ &= t\langle c_i, h \rangle + (1-t)\langle c_i, h \rangle = \\ &= \langle c_i, h \rangle. \end{aligned}$$

This and Remark 4.26 (ii) immediately allows us to conclude that

$$\widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{M}}_{c_i} \quad \text{for } i = 0, 1 \quad \Longleftrightarrow \quad \widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{M}}_{c_0} \cap \widetilde{\mathcal{M}}_{c_1}.$$

□

**Remark 4.29.** It follows from the previous remarks and Proposition 4.27, that, in general, the action minimizing measures (and consequently the mather sets  $\widetilde{\mathcal{M}}_c$  or  $\widetilde{\mathcal{M}}^h$ ) are not necessarily ergodic. Recall that an invariant probability measure is said to be *ergodic*, if all invariant Borel sets have measure 0 or 1. These measures play a special role in the study of the dynamics of the system, therefore one could ask what are the ergodic action-minimizing measures. It is a well-known result from ergodic theory, that the ergodic measures of a flow correspond to the *extremal points* of the set of invariant probability measures (see for instance [31]), where by “extremal point” of a convex set, we mean an element that cannot be obtained as a non-trivial convex combination of other elements of the set. Since  $\beta$  has superlinear growth, its epigraph  $\{(h, t) \in H_1(M; \mathbb{R}) \times \mathbb{R} : t \geq \beta(h)\}$  has infinitely many extremal points. Let  $(h, \beta(h))$  denote one of these extremal points. Then, there exists at least one ergodic action minimizing measure with rotation vector  $h$ . It is in fact sufficient to consider any extremal point of the set  $\{\mu \in \mathfrak{M}^h(L) : A_L(\mu) = \beta(h)\}$ : this measure will be an extremal point of  $\mathfrak{M}(L)$  and hence ergodic. Moreover, as we have already recalled in Remark 4.14, for such an ergodic measure  $\mu$ , Birkhoff's ergodic theorem implies that  $\mu$ -almost every trajectory of  $\Phi^L$  has rotation vector  $h$ .

## ADDENDA

### 4.A - THE SYMPLECTIC INVARIANCE OF MATHER SETS

In this addendum we would like to discuss some symplectic aspects of the theory that we have started to develop. In particular, as a first step, we would like to understand how the Mather sets behave under the action of *symplectomorphisms*

(see also [8]). In order to do this, we need to move to the Hamiltonian setting, rather than the Lagrangian one, and consider the Hamiltonian  $H : T^*M \rightarrow \mathbb{R}$ , defined by Fenchel duality (see Section 2). Recall that the associated Hamiltonian flow  $\Phi^H$  is conjugate to the Euler-Lagrange flow of  $L$ , therefore, from a dynamical systems point of view, the two systems are equivalent. We can define the associated Mather sets in the cotangent space as

$$\mathcal{M}_c^*(H) := \mathcal{L}_L(\widetilde{\mathcal{M}}_c(L)),$$

where  $\mathcal{L}_L : TM \rightarrow T^*M$  is the Legendre transform associated to  $L$  (see (4)). These sets are non-empty, compact and invariant for the Hamiltonian flow and they continue to satisfy the graph property (see Theorems 4.12 and 4.20).

The main advantage of shifting our point of view, is that the cotangent space  $T^*M$  can be naturally equipped with an exact symplectic form  $\omega = \sum_{i=1}^d dx_i \wedge dp_i = -d\lambda$ , where  $\lambda = \sum_{i=1}^d p_i dx_i$  is what is called the *tautological form* or *Liouville form*. Let  $\Psi : T^*M \rightarrow T^*M$  be a diffeomorphism. We shall say that  $\Psi$  is a *symplectomorphism* if it preserves the symplectic form  $\omega$ , i.e.,  $\Psi^*\omega = \omega$ . An easy class of examples is provided by *translations in the fibers*. Let  $\eta$  be a closed 1-form and consider  $\tau_\eta : T^*M \rightarrow T^*M$ ,  $(x, p) \mapsto (x, p + \eta(x))$ . Then,  $\tau$  is a symplectomorphism.

Observe that since  $\omega = -d\lambda$ , one has that  $d(\Psi^*\lambda - \lambda) = 0$ ; in other words, the 1-form  $\Psi^*\lambda - \lambda$  is closed. We shall call the *cohomology class* of  $\Psi$ , the cohomology class of  $\Psi^*\lambda - \lambda$  and denote it by  $[\Psi] \in H^1(T^*M; \mathbb{R}) \simeq H^1(M; \mathbb{R})$ , since  $T^*M$  can be retracted to  $M$  along the fibers (hereafter we shall always identify these two spaces). Obviously, going back to the previous example,  $[\tau_\eta] = [\eta]$ . In particular,  $\Psi$  is said to be *exact* if and only if  $[\Psi] = 0$ .

**Lemma 4.30.** *Any symplectomorphism  $\Psi : T^*M \rightarrow T^*M$  can be written as  $\Psi = \Phi \circ \tau_\eta$ , where  $\Phi$  is an exact symplectomorphism and  $[\eta] = [\Psi]$ .*

**Proof.** Let  $\eta$  be any closed 1-form on  $M$ , such that  $[\eta] = [\Psi]$ , and define  $\tau_\eta$  as above. Moreover, we define  $\Phi = \Psi \circ (\tau_\eta)^{-1} = \Psi \circ \tau_{-\eta}$ . In order to conclude the proof of the Lemma, we need to check that  $\Phi$  is exact. In fact,

$$\Phi^*\lambda = (\Psi \circ \tau_{-\eta})^*\lambda = \tau_{-\eta}^*(\Psi^*\lambda) = \tau_{-\eta}^*(\lambda + \Theta) = \lambda + \tau_{-\eta}^*\Theta - \eta,$$

where  $[\tau_{-\eta}^*\Theta] = [\Psi] = [\eta]$ . Then,  $[\Phi^*\lambda - \lambda] = [\Theta - \eta] = 0$  and hence it is exact.  $\square$

Therefore, if we want to understand the interplay between Mather sets and the action of symplectomorphisms, it will be sufficient to analyze the behaviour of these two kinds of symplectomorphisms: *translations in the fibers* and *exact symplectomorphisms*.

**Proposition 4.31.** *Let  $H : T^*M \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $\eta$  a closed 1-form on  $M$ . Then:*

$$\mathcal{M}_c^*(H \circ \tau_\eta) = \tau_{-\eta} \left( \mathcal{M}_{c+[\eta]}^*(H) \right) \quad \forall c \in H^1(M; \mathbb{R}).$$

**Remark 4.32.** If  $H$  is a Tonelli Hamiltonian, then also  $H \circ \tau_\eta$  is a Tonelli Hamiltonian (we are just composing it with a vertical translation in the fibers).

**Proof.** Let us start by observing that the Lagrangian associated to  $H \circ \tau_\eta$  is  $L_\eta := L - \hat{\eta}$  and that the associated Legendre transform  $\mathcal{L}_{L_\eta} = \tau_{-\eta} \circ \mathcal{L}$  (just derive  $L_\eta$  with respect to  $v$ ). Therefore, using the definition of Mather sets for a given cohomology class, we get:

$$\begin{aligned} \mathcal{M}_c^*(H \circ \tau) &:= \mathcal{L}_{L_\eta} \left( \widetilde{\mathcal{M}}_c(L - \hat{\eta}) \right) = \mathcal{L}_{L_\eta} \left( \widetilde{\mathcal{M}}_{c+[\eta]}(L) \right) = \\ &= (\tau_{-\eta} \circ \mathcal{L}_L) \left( \widetilde{\mathcal{M}}_{c+[\eta]}(L) \right) = \tau_{-\eta} \left( \mathcal{M}_{c+[\eta]}^*(H) \right). \end{aligned}$$

□

Let us see now what happens with exact symplectomorphisms (see also [54]).

**Proposition 4.33.** *Let  $H : T^*M \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $\Phi : T^*M \rightarrow T^*M$  an exact symplectomorphism such that  $H \circ \Phi$  is still of Tonelli type. Then:*

$$\mathcal{M}_c^*(H \circ \Phi) = \Phi^{-1}(\mathcal{M}_c^*(H)) \quad \forall c \in H^1(M; \mathbb{R}).$$

**Remark 4.34.** Observe that even if  $H$  is Tonelli,  $H \circ \Phi$  is not necessarily of Tonelli type!

This proposition can be easily deduced from the following Lemma (see also [54]).

**Lemma 4.35.** *Let  $H : T^*M \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $\Phi : T^*M \rightarrow T^*M$  an exact symplectomorphism, such that  $H' := H \circ \Phi$  is still Tonelli. Then:*

- (i)  $\mu$  is an invariant probability measure of  $H$  if and only if  $\Phi^*\mu$  is an invariant probability measure of  $H' = H \circ \Phi$ ;
- (ii) for any  $\mu$  invariant probability measure of  $H$ , the following holds:

$$(18) \quad \int \left[ p \frac{\partial H}{\partial p}(x, p) - H(x, p) \right] d\mu = \int \left[ p \frac{\partial H'}{\partial p}(x, p) - H'(x, p) \right] d\Phi_*\mu.$$

**Remark 4.36.** The identity in (18) represents the equality between the associated Lagrangian actions (in the Hamiltonian formalism). Therefore, since the respective actions coincide on all measures and the invariant measures are in 1-1 correspondence, there must be a correspondence between the minimizing ones:  $\mu$  is action minimizing for  $H$  if and only if  $\Phi^*\mu$  is action-minimizing for  $H \circ \Phi$ . From this, Proposition 4.33 follows easily.

**Proof. [Lemma 4.35]** (i) The first part of the statement follows from the classical fact that  $\Phi$  transforms the associated Hamiltonian vector fields in the following way:

$$(19) \quad X_H(\Phi(x, p)) = D\Phi(x, p)X_{H \circ \Phi}(x, p) \quad \forall (x, p) \in T^*M.$$

- (ii) If we denote by  $\lambda(x, p)$  the Liouville form  $pdx$ , then:

$$\int \left[ p \frac{\partial H}{\partial p}(x, p) - H(x, p) \right] d\mu = \int \left( \lambda(x, p)[X_H(x, p)] - H(x, p) \right) d\mu.$$

Therefore, using that  $|\det D\Phi| = 1$ , the relation in (19) and that  $\Phi^*\lambda - \lambda = df$  (since  $\Phi$  is an exact symplectomorphism), we obtain:

$$\begin{aligned} & \int \left[ p \frac{\partial H}{\partial p}(x, p) - H(x, p) \right] d\mu = \int \left( \lambda(x, p)[X_H(x, p)] - H(x, p) \right) d\mu = \\ &= \int \left( \Phi^*\lambda(x', p')[D\Phi^{-1}(x', p')X_H(\Phi(x', p'))] - H(\Phi(x', p')) \right) d\Phi^*\mu = \\ &= \int \left( (\lambda(x', p') + df(x', p'))[X_{H \circ \Phi}(x', p')] - H(\Phi(x', p')) \right) d\Phi^*\mu = \\ &= \int \left( \lambda(x', p')[X_{H'}(x', p')] - H'(x', p') \right) d\Phi^*\mu + \int df(x', p')[X_{H'}(x', p')] d\Phi^*\mu = \\ &= \int \left( p' \frac{\partial H'}{\partial p'}(x', p') - H'(x', p') \right) d\Phi^*\mu. \end{aligned}$$

In the last equality we used that  $\int df(x', p')[X_{H'}(x', p')] d\Phi^*\mu = 0$ , as it follows easily from the invariance of  $\Phi^*\mu$ . For the sake of simplifying the notation, let us denote  $\nu := d\Phi^*\mu$  and assume that  $\nu$  is ergodic (otherwise consider each ergodic

component). Using the ergodic theorem and the compactness of the support of  $\nu$ , we obtain that for a generic point  $(x_0, y_0)$  in the support of  $\nu$ :

$$\begin{aligned} \int df(x', p')[X_{H'}(x', p')] d\nu &= \lim_{N \rightarrow +\infty} \frac{1}{N} \int_0^N df(\Phi_t^{H'}(x_0, p_0))[X_{H'}(\Phi_t^{H'}(x_0, p_0))] dt = \\ &= \lim_{N \rightarrow +\infty} \frac{f(\Phi_N^{H'}(x_0, p_0)) - f(x_0, p_0)}{N} = 0. \end{aligned}$$

□

Finally, we can deduce the main result of this addendum.

**Theorem 4.37.** *Let  $H : T^*M \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $\Psi : T^*M \rightarrow T^*M$  a symplectomorphism of class  $[\Psi]$ , such that  $H \circ \Psi$  is of Tonelli type. Then,*

$$\mathcal{M}_c^*(H \circ \Psi) = \Psi^{-1} \left( \mathcal{M}_{c+[\Psi]}^*(H) \right) \quad \forall c \in H^1(M; \mathbb{R}).$$

**Proof.** Suppose that  $\Psi = \Phi \circ \tau_\eta$ , with  $[\eta] = [\Psi]$ . Then, using Propositions 4.31 and 4.33, we get for all  $c \in H^1(M; \mathbb{R})$ :

$$\begin{aligned} \mathcal{M}_c^*(H \circ \Psi) &= \mathcal{M}_c^*(H \circ \Phi \circ \tau_\eta) = \tau_{-\eta} \left( \mathcal{M}_{c+[\eta]}^*(H \circ \Phi) \right) = \\ &= \tau_\eta^{-1} \left( \Phi^{-1} \left( \mathcal{M}_{c+[\eta]}^*(H) \right) \right) = (\Phi \circ \tau_\eta)^{-1} \left( \mathcal{M}_{c+[\eta]}^*(H) \right) = \\ &= \Psi^{-1} \left( \mathcal{M}_{c+[\Psi]}^*(H) \right). \end{aligned}$$

□

**Corollary 4.38.** *Let  $H : T^*M \rightarrow \mathbb{R}$  be a Tonelli Hamiltonian and  $\Psi : T^*M \rightarrow T^*M$  a symplectomorphism of class  $[\Psi]$ , such that  $H \circ \Psi$  is of Tonelli type. Then:*

$$\begin{aligned} \text{(i)} \quad & \alpha_{H \circ \Psi}(c) = \alpha_H(c + [\Psi]) \quad \forall c \in H^1(M; \mathbb{R}) \\ \text{(ii)} \quad & \beta_{H \circ \Psi}(h) = \beta_H(h) - \langle [\Psi], h \rangle \quad \forall h \in H_1(M; \mathbb{R}). \end{aligned}$$

**Proof.** (i) From Theorem 4.37 it follows that if  $(x, p) \in \mathcal{M}^*(H \circ \Psi)$ , then  $\Psi(x, p) \in \mathcal{M}_{c+[\Psi]}^*(H)$ . Therefore, using Theorem 4.13, for any  $(x, p) \in \mathcal{M}^*(H \circ \Psi)$  we have:

$$-\alpha_{H \circ \Psi}(c) = (H \circ \Psi)(x, p) = H(\Psi(x, p)) = -\alpha_H(c + [\Psi]).$$

(ii) Observe that if  $h \in \partial \alpha_{H \circ \Psi}(c)$ , then  $h \in \partial \alpha_H(c + [\Psi])$ . In fact, for every  $c' \in H^1(M; \mathbb{R})$ :

$$\begin{aligned} \alpha_H(c' + [\Psi]) - \alpha_H(c + [\Psi]) &= \alpha_{H \circ \Psi}(c') - \alpha_{H \circ \Psi}(c) \geq \langle c' - c, h \rangle = \\ &= \langle (c' + [\Psi]) - (c + [\Psi]), h \rangle. \end{aligned}$$

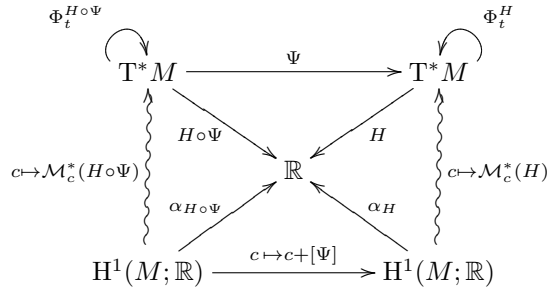
Let now  $h \in \partial \alpha_{H \circ \Psi}(c)$ . Using the fact that  $\alpha$  and  $\beta$  are one the conjugate of the other, we get:

$$\begin{aligned} \beta_{H \circ \Psi}(h) &= \langle c, h \rangle - \alpha_{H \circ \Psi}(c) = \langle c, h \rangle - \alpha_H(c + [\Psi]) = \\ &= \langle c + [\Psi], h \rangle - \alpha_H(c + [\Psi]) - \langle [\Psi], h \rangle = \\ &= \beta_H(h) - \langle [\Psi], h \rangle. \end{aligned}$$

□

We can summarize everything in the following commutative diagram.





#### 4.B - AN EXAMPLE: THE SIMPLE PENDULUM I

In this addendum we would like to describe the Mather sets, the  $\alpha$ -function and the  $\beta$ -function, in a specific example: the *simple pendulum*. This system can be described in terms of the Lagrangian:

$$\begin{aligned} L : \mathbb{T}\mathbb{T} &\longrightarrow \mathbb{R} \\ (x, v) &\longmapsto \frac{1}{2}|v|^2 + (1 - \cos(2\pi x)). \end{aligned}$$

It is easy to check that the Euler-Lagrange equation provides exactly the equation of the pendulum:

$$\dot{v} = 2\pi \sin(2\pi x) \quad \Longleftrightarrow \quad \begin{cases} v = \dot{x} \\ \ddot{x} - 2\pi \sin(2\pi x) = 0. \end{cases}$$

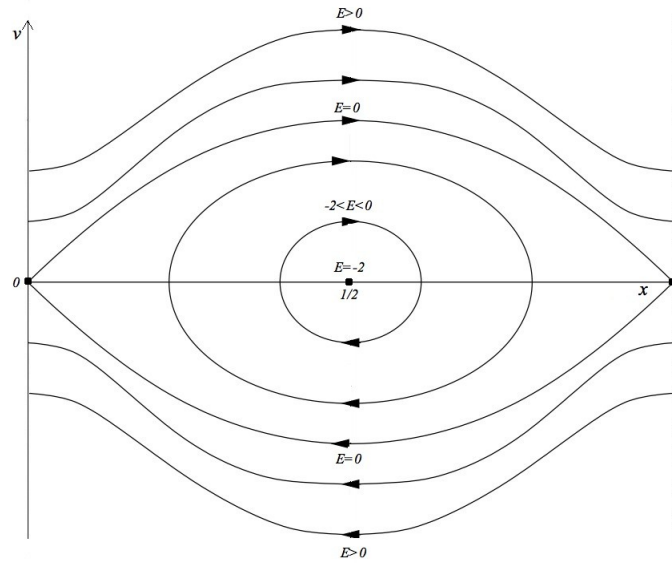


FIGURE 1. The phase space of the simple pendulum.

The associated Hamiltonian (or energy)  $H : T^*\mathbb{T} \rightarrow \mathbb{R}$  is given by  $H(x, p) := \frac{1}{2}|p|^2 - (1 - \cos(2\pi x))$ . Observe that in this case the Legendre transform  $\mathcal{L}_L(x, v) = (x, p)$ , therefore we can easily identify the tangent and cotangent space. In the following we shall consider  $T\mathbb{T} \simeq T^*\mathbb{T} \simeq \mathbb{T} \times \mathbb{R}$  and identify  $H^1(M; \mathbb{R}) \simeq H_1(M; \mathbb{R}) \simeq \mathbb{R}$ .

First of all, let us study what are the invariant probability measures of this system.

- Observe that  $(0, 0)$  and  $(\frac{1}{2}, 0)$  are fixed points for the system (respectively *unstable* and *stable*). Therefore, the Dirac measures concentrated on each of them are invariant probability measures. Hence, we have found two first invariant measures:  $\delta_{(0,0)}$  and  $\delta_{(\frac{1}{2},0)}$ , both with zero rotation vector:  $\rho(\delta_{(0,0)}) = \rho(\delta_{(\frac{1}{2},0)}) = 0$ . As far as their energy is concerned (*i.e.*, the energy levels in which they are contained), it is easy to check that  $E(\delta_{(0,0)}) = H(0, 0) = 0$  and  $E(\delta_{(\frac{1}{2},0)}) = H(\frac{1}{2}, 0) = -2$ . Observe that these two energy levels cannot contain any other invariant probability measure.
- If  $E > 0$ , then the energy level  $\{H(x, v) = E\}$  consists of two homotopically non-trivial periodic orbits (*rotation motions*):

$$\mathcal{P}_E^\pm := \{(x, v) : v = \pm \sqrt{2[(1 + E) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\}.$$

The probability measures evenly distributed along these orbits - which we shall denote  $\mu_E^\pm$  - are invariant probability measures of the system. If we denote by

$$(20) \quad T(E) := \int_0^1 \frac{1}{\sqrt{2[(1 + E) - \cos(2\pi x)]}} dx$$

the period of such orbits, then it is easy to check that (see Remark 4.14)  $\rho(\mu_E^\pm) = \frac{\pm 1}{T(E)}$ . Observe that this function  $T : (0, +\infty) \rightarrow (0, +\infty)$ , which associates to a positive energy  $E$  the period of the corresponding periodic orbits  $\mathcal{P}_E^\pm$ , is continuous and strictly decreasing. Moreover,  $T(E) \rightarrow \infty$  as  $E \rightarrow 0$  (it is easy to see this, noticing that motions on the *separatrices* take an “infinite” time to connect  $0$  to  $1 \equiv 0 \pmod{1}$ ). Therefore,  $\rho(\mu_E^\pm) \rightarrow 0$  as  $E \rightarrow 0$ .

- If  $-2 < E < 0$ , then the energy level  $\{H(x, v) = E\}$  consists of one contractible periodic orbit (*libration motion*):

$$\mathcal{P}_E := \{(x, v) : v^2 = 2(1 + E) - 2\cos(2\pi x), \quad x \in [x_E, 1 - x_E]\},$$

where  $x_E := \frac{1}{2\pi} \arccos(1 + E)$ . The probability measure evenly distributed along this orbit - which we shall denote  $\mu_E$  - is an invariant probability measure of the system. Moreover, since this orbit is contractible, its rotation vector is zero:  $\rho(\mu_E) = 0$ .

The measures above are the only ergodic invariant probability measures of the system. Other invariant measures can be easily obtained as convex combination of them.

Now we want to understand which of these are action-minimizing for some cohomology class.

**Remark 4.39.** (i) Let us start by remarking that for  $-2 < E < 0$  the support of the measure  $\mu_E$  is not a graph over  $\mathbb{T}$ , therefore it cannot be action-minimizing for any cohomology class, since otherwise it would violate Mather’s graph theorem

(Theorems 4.12 and 4.20). Therefore all action-minimizing measures will be contained in energy levels corresponding to energy bigger than zero. It follows from Theorem 4.13, that  $\alpha(c) \geq 0$  for all  $c \in \mathbb{R}$ .

(ii) Another interesting property of the  $\alpha$  function (in this specific case) is that it is an *even function*:  $\alpha(c) = \alpha(-c)$  for all  $c \in \mathbb{R}$ . This is a consequence of the particular symmetry of the system, *i.e.*,  $L(x, v) = L(x, -v)$ . In fact, let us denote  $\tau : \mathbb{T} \times \mathbb{R} \longrightarrow \mathbb{T} \times \mathbb{R}$ ,  $(x, v) \longmapsto (x, -v)$  and observe that if  $\mu$  is an invariant probability measure, then also  $\tau^*\mu$  is still an invariant probability measure. Moreover,  $\tau^*\mathfrak{M}(L) = \mathfrak{M}(L)$ , where  $\mathfrak{M}(L)$  denotes the set of all invariant probability measures of  $L$ . It is now sufficient to notice that for each  $\mu \in \mathfrak{M}(L)$ ,  $\int (L - c \cdot v) d\mu = \int (L + c \cdot v) d\tau^*\mu$ , and hence conclude that

$$\alpha(c) = - \inf_{\mathfrak{M}(L)} \int (L - c \cdot v) d\mu = - \inf_{\mathfrak{M}(L)} \int (L + c \cdot v) d\tau^*\mu = \alpha(-c).$$

(iii) It follows from the above symmetry and the convexity of  $\alpha$ , that  $\min_{\mathbb{R}} \alpha(c) = \alpha(0)$ .

Let us now start by studying the 0-action minimizing measures, *i.e.*, invariant probability measures that minimize the action of  $L$  without any “correction”. Since  $L(x, v) \geq 0$  for each  $(x, v) \in \mathbb{T} \times \mathbb{R}$ , then  $A_L(\mu) \geq 0$  for all  $\mu \in \mathfrak{M}(L)$ . In particular,  $A_L(\delta_{(0,0)}) = 0$ , therefore  $\delta_{(0,0)}$  is a 0-action minimizing measure and  $\alpha(0) = 0$ . Since there are not other invariant probability measures supported in the energy level  $\{H(x, v) = 0\}$  (*i.e.*, on the separatrices), then we can conclude that:

$$\widetilde{\mathcal{M}}_0 = \{(0, 0)\}.$$

Moreover, since  $\alpha'(0) = 0$  (see Remark 4.39 (iii)), then it follows from Remark 4.26 that:

$$\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0 = \{(0, 0)\}.$$

On the other hand, this could be also deduced from the fact that the only other measures with rotation vector 0, cannot be action minimizing since they do not satisfy the graph theorem (Remark 4.39).

Now let us investigate what happens for other cohomology classes. A naïve observation is that since the  $\alpha$  function is superlinear and continuous, all energy levels for  $E > 0$  must contain some Mather set; in other words, they will be achieved for some  $c$ .

Let  $E > 0$  and consider the periodic orbit  $\mathcal{P}_E^+$  and the invariant probability measure  $\mu_E^+$  evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form  $\eta_E^+ := \sqrt{2[(1+E) - \cos(2\pi x)]} dx$ , whose cohomology class is

$$(21) \quad c^+(E) := [\eta_E^+] = \int_0^1 \sqrt{2[(1+E) - \cos(2\pi x)]} dx,$$

which can be interpreted as the (signed) area between the curve and the positive  $x$ -semiaxis. This value is clearly continuous and strictly increasing with respect to  $E$  (for  $E > 0$ ) and, as  $E \rightarrow 0$ :

$$c^+(E) \longrightarrow \int_0^1 \sqrt{2[1 - \cos(2\pi x)]} dx = \frac{4}{\pi}.$$

Therefore, it defines an invertible function  $c^+ : (0, +\infty) \longrightarrow (\frac{4}{\pi}, +\infty)$ .

We want to prove that  $\mu_E^+$  is  $c^+(E)$ -action minimizing. The proof will be an imitation of what already seen for KAM tori in Section 3 (see Proposition 3.4).

Let us consider the Lagrangian  $L_{\eta_E^+}(x, v) := L(x, v) - \eta_E^+(x) \cdot v$ . Then, using Fenchel-Legendre inequality (3) (on the support of  $\mu_E^+$ , because of our choice of  $\eta_E^+$ , this is indeed an equality):

$$\begin{aligned} \int L_{\eta_E^+}(x, v) d\mu_E^+ &= \int (L(x, v) - \eta_E^+(x) \cdot v) d\mu_E^+ = \\ &= \int -H(x, \eta_E^+(x)) d\mu_E^+ = -E. \end{aligned}$$

Now, let  $\nu$  be any other invariant probability measure and apply again the same procedure as above (warning: this time Fenchel-Legendre inequality is not an equality anymore!):

$$\begin{aligned} \int L_{\eta_E^+}(x, v) d\nu &= \int (L(x, v) - \eta_E^+(x) \cdot v) d\nu \geq \\ &\geq \int -H(x, \eta_E^+(x)) d\nu = -E. \end{aligned}$$

Therefore, we can conclude that  $\mu_E^+$  is  $c^+(E)$ -action minimizing. Since it already projects over the whole  $\mathbb{T}$ , it follows from the graph theorem that it is the only one:

$$\widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+ = \{(x, v) : v = \sqrt{2[(1+E) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\}.$$

Furthermore, since  $\rho(\mu_E^+) = \frac{1}{T(E)}$ , then:

$$\widetilde{\mathcal{M}}_{\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+.$$

Similarly, one can consider the periodic orbit  $\mathcal{P}_E^-$  and the invariant probability measure  $\mu_E^-$  evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form  $\eta_E^- := -\sqrt{2[(1+E) - \cos(2\pi x)]} dx = -\eta_E^+$ , whose cohomology class is  $c^-(E) = -c^+(E)$ . Then (see also Remark 4.39 (ii)):

$$\widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^- = \{(x, v) : v = -\sqrt{2[(1+E) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\},$$

and

$$\widetilde{\mathcal{M}}_{-\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^-.$$

Note that this completes the study of the Mather sets for any given rotation vector, since

$$\rho(\mu_E^\pm) = \pm \frac{1}{T(E)} \xrightarrow{E \rightarrow +\infty} \pm \infty \quad \text{and} \quad \rho(\mu_E^\pm) = \pm \frac{1}{T(E)} \xrightarrow{E \rightarrow 0^+} 0.$$

What remains to study is what happens for non-zero cohomology classes in  $[-\frac{4}{\pi}, \frac{4}{\pi}]$ . The situation turns out to be quite easy. Observe that  $\alpha(c^\pm(E)) = E$ . Therefore, from the continuity of  $\alpha$  it follows that (take the limit as  $E \rightarrow 0$ ):  $\alpha(\pm\frac{4}{\pi}) = 0$ . Moreover, since  $\alpha$  is convex and  $\min \alpha(c) = \alpha(0) = 0$ , then:  $\alpha(c) \equiv 0$  on  $[-\frac{4}{\pi}, \frac{4}{\pi}]$ . Therefore, the corresponding Mather sets will lie in the zero energy level. From the above discussion, it follows that in this energy level there is a unique invariant probability measure, namely  $\delta_{(0,0)}$ , and consequently:

$$\widetilde{\mathcal{M}}_c = \{(0, 0)\} \quad \text{for all } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi}.$$

Let us summarize what we have found so far. Recall that in (20) and (21) we have introduced these two functions:  $T : (0, +\infty) \rightarrow (0, +\infty)$  and  $c^+ : (0, +\infty) \rightarrow (\frac{4}{\pi}, +\infty)$  representing respectively the period and the “cohomology” (area below the curve) of the “upper” periodic orbit of energy  $E$ . These functions (for which we have an explicit formula in terms of  $E$ ) are continuous and strictly monotone (respectively, decreasing and increasing). Therefore, we can define their inverses

which provide the energy of the periodic orbit with period  $T$  (for all positive periods) or the energy of the periodic orbit with cohomology class  $c$  (for  $|c| > \frac{4}{\pi}$ ). We shall denote them  $E(T)$  and  $E(c)$  (observe that this last quantity is exactly the  $-\alpha(c)$ ). Then:

$$\widetilde{\mathcal{M}}_c = \begin{cases} \{(0,0)\} & \text{if } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi} \\ \mathcal{P}_{E(c)}^+ & \text{if } c > \frac{4}{\pi} \\ \mathcal{P}_{E(-c)}^- & \text{if } c < -\frac{4}{\pi} \end{cases}$$

and

$$\widetilde{\mathcal{M}}^h = \begin{cases} \{(0,0)\} & \text{if } h = 0 \\ \mathcal{P}_{E(\frac{1}{h})}^+ & \text{if } h > 0 \\ \mathcal{P}_{E(-\frac{1}{h})}^- & \text{if } h < 0. \end{cases}$$

We can provide an expression for these functions in terms of the quantities introduced above:

$$\alpha(c) = \begin{cases} 0 & \text{if } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi} \\ E(|c|) & \text{if } |c| > \frac{4}{\pi} \end{cases}$$

and

$$\beta(h) = \begin{cases} 0 & \text{if } h = 0 \\ c(E(\frac{1}{|h|}))|h| - E(\frac{1}{|h|}) & \text{if } h \neq 0. \end{cases}$$

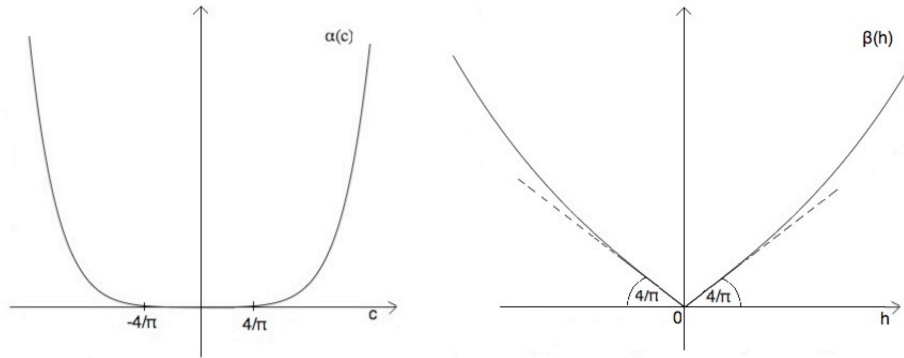


FIGURE 2. Sketch of the graphs of the  $\alpha$  and  $\beta$  functions of the simple pendulum.

Observe that the  $\alpha$  function is  $C^1$ . In fact, the only problem might be at  $c = \pm \frac{4}{\pi}$ , but also there it is differentiable, with derivative 0. If it were not differentiable, then there would exist a subderivative  $h \neq 0$  and consequently  $\widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_{\pm \frac{4}{\pi}}$ , which is absurd since the set on the right-hand side consists of a single point. However,  $\alpha$  is not strictly convex, since there is a *flat* piece on which it is zero.

As far as  $\beta$  is concerned, it is strictly convex (as a consequence of  $\alpha$  being  $C^1$ ), but it is differentiable everywhere except at the origin. At the origin, in fact, there is a corner and the set of subderivatives (*i.e.*, the slopes of tangent lines) is given by  $\partial\beta(0) = [-\frac{4}{\pi}, \frac{4}{\pi}]$  (this is related to the fact that  $\alpha$  has a flat on this interval).

#### 4.C - HOLONOMIC MEASURES AND GENERIC PROPERTIES

In this addendum we would like to stress that using the above approach the minimizing measures are obtained through a variational principle over the set of

invariant probability measures. Because of the request of “invariance”, this set clearly depends on the Lagrangian that one is considering. Moreover, it is somehow unnatural for variational problems to ask “a-priori invariance”. What generally happens, in fact, is that invariance is obtained as a byproduct of the minimization process carried out.

An alternative approach, slightly different under this respect, was due to Ricardo Mañé [33] (see also [34]). This deals with the bigger set of *holonomic measures* (or *closed measures*, see Remark 4.40) and prove extremely advantageous when dealing with different Lagrangians at the same time. In this addendum we want to sketch the basic ideas behind it.

Let  $C_\ell^0$  be the set of continuous functions  $f : TM \rightarrow \mathbb{R}$  growing (fiberwise) at most linearly, *i.e.*,

$$\|f\|_\ell := \sup_{(x,v) \in TM} \frac{f(x,v)}{1 + \|v\|} < +\infty,$$

and let  $\mathfrak{M}_M^\ell$  be the set of probability measures on the Borel  $\sigma$ -algebra of  $TM$  such that  $\int_{TM} \|v\| d\mu < \infty$ , endowed with the unique metrizable topology given by:

$$\mu_n \longrightarrow \mu \quad \Longleftrightarrow \quad \int_{TM} f(x,v) d\mu_n \longrightarrow \int_{TM} f(x,v) d\mu \quad \forall f \in C_\ell^0.$$

Let  $(C_\ell^0)^*$  be the dual of  $C_\ell^0$ . Then  $\mathfrak{M}_M^\ell$  can be naturally embedded in  $(C_\ell^0)^*$  and its topology coincides with that induced by the weak\* topology on  $(C_\ell^0)^*$ . One can show that this topology is metrizable and a metric is, for instance:

$$d(\mu_1, \mu_2) = \left| \int_{TM} \|v\| d\mu_1 - \int_{TM} \|v\| d\mu_2 \right| + \sum_n \frac{1}{2^n c_n} \left| \int_{TM} \varphi_n d\mu_1 - \int_{TM} \varphi_n d\mu_2 \right|,$$

where  $\{\varphi_n\}_n$  is a sequence of functions with compact support on  $C_\ell^0$ , which is dense on  $C_\ell^0$  (in the topology of uniform convergence on compact subsets of  $TM$ ) and  $c_n := \sup_{TM} |\varphi_n(x,v)|$ . The space of probability measures that we shall be considering is a closed subset of  $\mathfrak{M}_M^\ell$  (endowed with the induced topology), which is defined as follows. If  $\gamma : [0, T] \rightarrow M$  is a closed absolutely continuous curve, let  $\mu_\gamma$  be such that

$$\int_{TM} f(x,v) d\mu_\gamma = \frac{1}{T} \int_0^T f(\gamma(t), \dot{\gamma}(t)) dt \quad \forall f \in C_\ell^0.$$

Observe that  $\mu_\gamma \in \mathfrak{M}_M^\ell$  because if  $\gamma$  is absolutely continuous then  $\int |\dot{\gamma}(t)| dt < +\infty$ . Let  $\mathcal{C}(M)$  be the set of such  $\mu_\gamma$ 's and  $\overline{\mathcal{C}(M)}$  its closure in  $\mathfrak{M}_M^\ell$ . This set is convex and it is called the set of *holonomic measures* on  $M$ .

One can check that the following properties are satisfied (see [34]):

- i)  $\mathfrak{M}(L) \subseteq \overline{\mathcal{C}(M)} \subseteq \mathfrak{M}_M^\ell$ . In particular, for every Tonelli Lagrangian  $L$  on  $TM$ , all probabilities measures  $\mu$  that are invariant with respect to the Euler-Lagrange flow and such that  $\int_{TM} L d\mu < +\infty$ , are contained in  $\overline{\mathcal{C}(M)}$ .
- ii) To any given probability  $\mu \in \mathcal{C}(M)$ , one can associate a rotation vector  $\rho(\mu) \in H_1(M; \mathbb{R})$ . This map extends continuously to a map

$$\rho : \overline{\mathcal{C}(M)} \longrightarrow H_1(M; \mathbb{R})$$

and this extension is surjective.

- iii) For each  $C \in \mathbb{R}$  the set  $\left\{ \mu \in \overline{\mathcal{C}(M)} : A_L(\mu) \leq C \right\}$  is compact.
- iv) If a measure  $\mu \in \overline{\mathcal{C}(M)}$  satisfies

$$A_L(\mu) = \min \left\{ A_L(\nu) : \nu \in \overline{\mathcal{C}(M)} \right\},$$

then  $\mu \in \mathfrak{M}(L)$  (and in particular it is invariant). Observe that the existence of probabilities attaining the minimum follows from iii).

In view of these properties, it is clear that the corresponding minimizing problem, although on a bigger space of measures, will lead to the same results as before and the same definition of Mather sets.

**Remark 4.40.** One can define the set of *closed measures* on  $TM$  as:

$$\mathcal{K}(TM) := \left\{ \mu \in \mathfrak{M}_M^\ell \text{ such that } \int df(x) \cdot v d\mu = 0 \ \forall f \in C^1(M) \right\}.$$

It is easy to verify that holonomic measures satisfy Proposition 4.7 and therefore  $\overline{\mathcal{C}(M)} \subseteq \mathcal{K}(TM)$ . It is definitely less trivial to prove that indeed these two sets coincide:  $\overline{\mathcal{C}(M)} \subseteq \mathcal{K}(TM)$ . Although this was originally noticed by John Mather, as far as I know it has never been published by the author. A proof of this result may be found in [10].

As we have already pointed out, this different approach is more suitable for working with different Lagrangians, for instance if one wants to study properties of family of Lagrangians or want to do some sort of perturbation theory. Using these ideas, Ricardo Mañé [34] showed that one can prove much stronger results if, instead of considering ALL Lagrangians, one considers *generic Lagrangians*.

**Definition 4.41.** A property  $P$  is said to be *generic* (in the sense of Mañé) for a Lagrangian  $L$  if there exists a residual set (i.e., dense  $G_\delta$  set)  $\mathcal{S}_L \subseteq C^2(M)$ , such that if  $U \in \mathcal{S}_L$  then  $L + U$  satisfies the property  $P$ .

**Theorem 4.42 (Mañé).** For a fixed  $c \in H^1(M; \mathbb{R})$ , having a unique  $c$ -action minimizing measure is a generic property in the sense of Mañé. In other words, for any Tonelli Lagrangian  $L$ , there exists a residual subset  $\mathcal{S}_L \subseteq C^2(M)$ , such that for each  $U \in \mathcal{S}_L$ , the Lagrangian  $L + U$  has a unique  $c$ -action minimizing measure.

This result has been recently improved by Patrick Bernard and Gonzalo Contreras [11]:

**Theorem 4.43 (Bernard-Contreras).** The following property is generic in the sense of Mañé: for all  $c \in H^1(M; \mathbb{R})$ , there are at most  $1 + \dim H^1(M; \mathbb{R})$  ergodic  $c$ -action minimizing measures.

## 5. GLOBAL ACTION MINIMIZING CURVES: AUBRY AND MAÑÉ SETS

In the previous section we have described the construction and the main properties of the Mather sets. One of the main limitations of these sets is that, being the support of invariant probability measures, they are *recurrent* under the flow (Poincaré recurrence theorem), i.e., each orbit after a sufficiently long time (and therefore infinitely many often) will return arbitrarily close to its initial point. This property excludes many interesting invariant sets, which are somehow “invisible” to such a construction; for instance, think about the stable and unstable manifolds of some hyperbolic invariant set, or about heteroclinic and homoclinic orbits between invariant sets.

In this section we shall construct other (possibly) “larger” compact invariant sets and discuss their significance for the dynamics: the *Aubry sets* and the *Mañé sets*. The key idea is the same as we have already explored in Section 3: instead of considering action minimizing invariant probability measures, one can look at *action-minimizing curves* for some modified Lagrangian. We showed in Section 3 (see Remarks 3.9 and 3.12) that orbits on KAM tori could be characterized in terms of

this property. In this section we shall mimic that construction in the general case of a Tonelli Lagrangians.

In the light of Lemma 4.6 and the discussion in Section 4, let us fix a cohomology class  $c \in H^1(M; \mathbb{R})$  and choose a smooth 1-form  $\eta$  on  $M$  that represents  $c$ . As we have already pointed out in Section 2, there is a close relation between solutions of the Euler-Lagrange flow and extremals of the action functional  $A_{L_\eta}$  for the fixed end-point problem (which are the same as the extremals of  $A_L$ ). In general, these extremals are not minima (they are local minima only if the time length is very short [22, Section 3.6]). One could wonder if such minima exist, namely if for any given end-points  $x, y \in M$  and any given positive time  $T$ , there exists a minimizing curve connecting  $x$  to  $y$  in time  $T$ . From what already said, this curve will correspond to an orbit for the Euler-Lagrange flow. Under our hypothesis on the Lagrangian, the answer to this question turns out to be affirmative. This is a classical result in calculus of variations, known as *Tonelli Theorem*.

**Theorem 5.1 (Tonelli Theorem, [42]).** *Let  $M$  be a compact manifold and  $L$  a Tonelli Lagrangian on  $TM$ . For all  $a < b \in \mathbb{R}$  and  $x, y \in M$ , there exists, in the set of absolutely continuous curves  $\gamma : [a, b] \rightarrow M$  such that  $\gamma(a) = x$  and  $\gamma(b) = y$ , a curve that minimizes the action  $A_{L_\eta}(\gamma) = \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt$ .*

**Remark 5.2.** (i) A curve minimizing  $A_{L_\eta}(\gamma) = \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt$  subject to the fixed end-point condition  $\gamma(a) = x$  and  $\gamma(b) = y$ , is called a *c-Tonelli minimizer*. Recall that such minimizers do only depend on  $c$  and not on the chosen representative  $\eta$ . In fact, adding an exact 1-form  $df$  to  $L$  will contribute with a constant term  $f(y) - f(x)$ , that does not play any role in selecting the minimizers (see also Proposition 4.7).

(ii) As Mañé pointed out in [35], for these minimizers to exist it is not necessary to assume the compactness of  $M$ : the superlinear growth condition with respect to some complete Riemannian metric on  $M$  is enough.

(iii) A Tonelli minimizer which is  $C^1$  is in fact  $C^r$  (if the Lagrangian  $L$  is  $C^r$ ) and satisfies the Euler-Lagrange equation; this follows from the usual elementary arguments in calculus of variations, together with Caratheodory's remark on differentiability. In the autonomous case, Tonelli minimizers will be always  $C^1$ . In the non-autonomous time-periodic case (Tonelli Theorem holds also in this case [42]), as already remarked in remark 2.2, one needs to require that the Euler-Lagrange flow is also complete.

We shall sketch here the proof of Tonelli Theorem. We refer the reader to [42, Appendix 1] for more details. A new shorter proof of this theorem can be found in [10].

**Proof. [Tonelli Theorem]** The proof of this theorem follows from the following result (by  $C^{ac}$  we denote the set of absolutely continuous functions).

*Lemma.* *Let  $K \in \mathbb{R}$ . The set  $S_K := \{\gamma \in C^{ac}([a, b], M) : A_{L_\eta}(\gamma) \leq K\}$  is compact in the  $C^0$  topology.*

First let us see how to obtain Tonelli Theorem from this Lemma. Let  $k_0 := \inf\{A_{L_\eta}(\gamma) : \gamma \in C^{ac}([a, b], M)\}$ . Observe that  $k_0 > -\infty$  since  $L_\eta$  is bounded from below. Therefore, for any  $K > k_0$ ,  $S_K \neq \emptyset$  and it is compact because of the Lemma. Moreover,  $S_K \subseteq S_{K'}$  if  $K \leq K'$ . Hence:

$$\bigcap_{K > k_0} S_K \neq \emptyset$$



and any element in this intersection is a  $c$ -Tonelli minimizer.

The proof of the above lemma consists of several steps (see [42, Appendix 1] for the missing details).

- The first step is the observation that the family of curves in  $S_k$  is absolutely equicontinuous, *i.e.*, for every  $\varepsilon > 0$  there exist  $\delta > 0$  such that if  $a \leq a_0 < b_0 \leq a_1 < b_1 \leq \dots \leq a_n < b_n \leq b$  and  $\sum_{i=0}^n (b_i - a_i) < \delta$ , then  $\sum_{i=0}^n d(\gamma(a_i), \gamma(b_i)) < \varepsilon$ . For this, one needs to use the superlinearity of  $L$ .
- Now, we can apply Ascoli-Arzelà theorem to deduce that every sequence  $\{\gamma_n\}_n$  in  $S_K$  has a convergent subsequence, with respect to the  $C^0$  topology. Moreover, it follows easily from the definition of absolute equicontinuity, that the limit of any convergent subsequence must be also absolutely equicontinuous.
- The last, and more involved, step is to show that if  $\gamma$  is the limit of a sequence  $\{\gamma_n\}_n$  in  $S_K$ , then  $\gamma \in S_k$ , namely  $A_{L_\eta}(\gamma) \leq K$ . See [42, pages 199-201].

These three steps conclude the proof of the lemma.  $\square$

**Remark 5.3.** Observe that the lemma used in the proof of Tonelli theorem is a sort of semicontinuity result for the Lagrangian action (compare with Proposition 4.3). In fact, it implies that if  $\{\gamma_n\}_n$  is a sequence in  $C^{ac}([a, b], M)$  which converges to  $\gamma$  in the  $C^0$  topology, then  $\gamma \in C^{ac}([a, b], M)$  and  $A_{L_\eta}(\gamma) \leq \liminf_{n \rightarrow +\infty} A_{L_\eta}(\gamma_n)$ .

In the following we shall be interested in particular Tonelli minimizers that are defined for all times and whose action is minimal with respect to any given time length. We shall see that these curves present a very rich structure.

**Definition 5.4 (c-minimizers).** *An absolutely continuous curve  $\gamma : \mathbb{R} \rightarrow M$  is a  $c$ -(global) minimizer for  $L$ , if for any given  $a < b \in \mathbb{R}$*

$$A_{L_\eta}(\gamma|[a, b]) = \min A_{L_\eta}(\sigma)$$

where the minimum is taken over all  $\sigma : [a, b] \rightarrow M$  such that  $\sigma(a) = \gamma(a)$  and  $\sigma(b) = \gamma(b)$ .

We have already seen in Section 3 that one can give an a-priori stronger notion of minimizer, asking that the minimum is realized amongst all curves connecting the two end-points, independently of their time length.

**Definition 5.5 (c-time free minimizers).** *An absolutely continuous curve  $\gamma : \mathbb{R} \rightarrow M$  is a  $c$ -time free minimizer for  $L$ , if for any given  $a < b \in \mathbb{R}$*

$$A_{L_\eta}(\gamma|[a, b]) = \min A_{L_\eta}(\sigma)$$

where the minimum is taken over all  $\sigma : [a', b'] \rightarrow M$  such that  $\sigma(a') = \gamma(a)$  and  $\sigma(b') = \gamma(b)$ .

**Remark 5.6.** (i) We have proved in section 3 that orbits on KAM tori satisfy this stronger condition, modulo adding a constant to the Lagrangian (see Remark 3.9). In fact, it is quite easy to see that this condition is “sensitive” to the addition of constants to the Lagrangian (although this is something totally irrelevant for being or not a  $c$ -minimizer). For example, suppose that  $\int_0^T L_\eta(\gamma, \dot{\gamma}) dt < \int_0^{T'} L_\eta(\sigma, \dot{\sigma}) dt$  with  $T' < T$  and let  $k$  be a constant such that  $k > \frac{1}{T' - T} \left( \int_0^{T'} L_\eta(\sigma, \dot{\sigma}) dt - \int_0^T L_\eta(\gamma, \dot{\gamma}) dt \right)$ .

Then, adding  $k$  to the Lagrangian, we even reverse the inequality:

$$\begin{aligned}
A_{L_\eta+k}(\gamma) &= \int_0^T L_\eta(\gamma, \dot{\gamma}) dt + kT = \int_0^T L_\eta(\gamma, \dot{\gamma}) dt + k(T - T') + kT' \geq \\
&\geq \int_0^T L_\eta(\gamma, \dot{\gamma}) dt + \int_0^{T'} L_\eta(\sigma, \dot{\sigma}) dt - \int_0^T L_\eta(\gamma, \dot{\gamma}) dt + kT' = \\
&= \int_0^{T'} L_\eta(\sigma, \dot{\sigma}) dt + kT' = A_{L_\eta+k}(\sigma).
\end{aligned}$$

(ii) Obviously a  $c$ -time free minimizer is also a  $c$ -minimizer. Fathi [20] proved that in the autonomous case, modulo adding the (unique) right constant, these two notions of minimizers indeed coincide: *if  $\gamma$  be a  $c$ -minimizer for  $L$ , then,  $\gamma$  is  $c$ -time free minimizer for  $L + \alpha(c)$ , where  $\alpha$  is the  $\alpha$ -function associated to  $L$ .* We proved this in the special setting of Section 3 (see Proposition 3.8). We shall discuss the general case in Section 6.

(iii) The equivalence between these two notions of minimizers is not true anymore when we consider time-periodic Tonelli Lagrangians (see [25]). Tonelli Lagrangians for which this equivalence result holds are called *regular*. Patrick Bernard [5] showed that under suitable assumptions on the Mather set it is possible to prove that the Lagrangian is regular. For instance, if the Mather set  $\mathcal{M}_c$  is union of 1-periodic orbits, then  $L_\eta$  is regular. This problem turned out to be strictly related to the convergence of the so-called *Lax-Oleinik semigroup* (see [22] and Section 6 for its definition).

Now we shall study the existence and the properties of  $c$ -minimizers and  $c$ -time free minimizers. There are two equivalent approaches that one can pursue: one is essentially due to Ricardo Mañé [35, 14], while the other has been developed by John Mather [44]. In the following, in order to keep the analogy with the cartoon example discussed in Section 3, we prefer following the first of these two approaches. We shall discuss Mather's approach in an Addendum 5C at the end of this section.

Given any  $x, y \in M$  and  $T > 0$ , let us denote by  $C_T(x, y)$  the set of absolutely continuous curves  $\gamma : [0, T] \rightarrow M$  such that  $\gamma(0) = x$  and  $\gamma(T) = y$ . Tonelli Theorem implies that there exists  $\gamma_{min} \in C_T(x, y)$  realizing the minimum, i.e.,  $A_{L_\eta}(\gamma_{min}) = \min_{\gamma \in C_T(x, y)} A_{L_\eta}(\gamma)$ . Our goal here is to study the existence of  $c$ -time free minimizers for  $L$ . For, let us fix  $k \in \mathbb{R}$  and consider the following quantity:

$$\phi_{\eta,k}(x, y) = \inf_{T>0} \min_{\gamma \in C_T(x, y)} A_{L_\eta+k}(\gamma) \in \mathbb{R} \cup \{-\infty\}.$$

This quantity is commonly called *Mañé potential* (compare with (9) in Section 3). First of all we would like to understand when it is finite and what are its properties. Let us introduce what is called *Mañé critical value*.

**Definition 5.7 (Mañé critical value).**

$$\begin{aligned}
c(L_\eta) &:= \sup\{k \in \mathbb{R} : \exists \text{ a closed curve } \gamma \text{ s.t. } A_{L_\eta+k}(\gamma) < 0\} = \\
&= \inf\{k \in \mathbb{R} : \forall \text{ closed curves } \gamma \text{ s.t. } A_{L_\eta+k}(\gamma) \geq 0\}.
\end{aligned}$$

**Remark 5.8.** (i) It is easy to check that  $c(L_\eta) < \infty$ . In fact, since  $L$  is superlinear, there exists a sufficiently large  $k$  such that  $L + k \geq 0$  everywhere.

(ii) Moreover  $c(L_\eta)$  only depends on  $c = [\eta]$  and not on the chosen representative. In fact, it is sufficient to notice that the integral of exact 1-forms along closed curves is zero. We shall see in the following that this “critical value” is something that we have already met before:  $c(L_\eta) = \alpha(c)$ , where  $\alpha$  is Mather's  $\alpha$ -function associated to  $L$ . In Section 6 we shall also point out its relation to viscosity solutions

and subsolutions of Hamilton-Jacobi equation and the critical value introduced by Lions, Papanicolau and Varadhan in [32].

**Proposition 5.9** (See also [14]).

- (1)  $\forall k \in \mathbb{R} : \quad \phi_{\eta,k}(x, y) \leq \phi_{\eta,k}(x, z) + \phi_{\eta,k}(z, y) \quad \forall x, y, z \in M.$
- (2)  $\text{If } k < c(L_\eta) : \quad \phi_{\eta,k}(x, y) \equiv -\infty \quad \forall x, y \in M.$   
 $\text{If } k \geq c(L_\eta) : \quad \phi_{\eta,k}(x, y) \in \mathbb{R} \quad \forall x, y \in M.$
- (3)  $\text{If } k \geq c(L_\eta) : \quad \phi_{\eta,k} : M \times M \longrightarrow \mathbb{R} \text{ is Lipschitz.}$
- (4)  $\text{If } k \geq c(L_\eta) : \quad \phi_{\eta,k}(x, x) \equiv 0 \quad \forall x \in M.$
- (5)  $\text{If } k \geq c(L_\eta) : \quad \phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) \geq 0 \quad \forall x, y \in M.$   
 $\text{If } k > c(L_\eta) : \quad \phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) > 0 \quad \forall x \neq y \in M.$

**Proof.** (1) First of all observe that this inequality makes sense also if  $\phi_{\eta,k}(x, y) = -\infty$  for some  $x, y \in M$ . Let  $\gamma_1 \in C_T(x, z)$  and  $\gamma_2 \in C_{T'}(z, y)$  and consider the new curve obtained by joining them:  $\gamma_1 * \gamma_2 \in C_{T+T'}(x, y)$ . Since the action is linear, it follows that  $A_{L_\eta}(\gamma_1 * \gamma_2) = A_{L_\eta}(\gamma_1) + A_{L_\eta}(\gamma_2)$  and therefore

$$\phi_{\eta,k}(x, y) \leq A_{L_\eta}(\gamma_1) + A_{L_\eta}(\gamma_2).$$

It is now sufficient to take the infimum over all possible  $(\gamma_1, T)$  and  $(\gamma_2, T')$  to conclude that

$$\phi_{\eta,k}(x, y) \leq \phi_{\eta,k}(x, z) + \phi_{\eta,k}(z, y).$$

(2) We shall first prove that if  $k < c(L_\eta)$  there exists  $x_0 \in M$  such that  $\phi_{\eta,k}(x_0, x_0) = -\infty$ . In fact, from the definition of  $c(L_\eta)$ , we know that there exists  $\gamma : [0, T] \longrightarrow M$  closed curve with  $A_{L_\eta}(\gamma) < 0$ . Let us denote by  $\gamma^n$  the  $n$ -time iteration of  $\gamma$ , i.e.,  $\gamma^n := \gamma * \dots * \gamma$  ( $n$  times). Then,

$$\phi_{\eta,k}(\gamma(0), \gamma(0)) \leq A_{L_\eta}(\gamma^n) = nA_{L_\eta}(\gamma) \xrightarrow{n \rightarrow \infty} -\infty.$$

Choose  $x_0 := \gamma(0)$  (or any other point on  $\gamma$ ). The first claim will now follow from (1). In fact, if  $x, y \in M$ , then:

$$\phi_{\eta,k}(x, y) \leq \phi_{\eta,k}(x, x_0) + \phi_{\eta,k}(x_0, x_0) + \phi_{\eta,k}(x_0, y) = -\infty.$$

As for the second claim, if  $k \geq c(L_\eta)$  it follows from the definition of  $c(L_\eta)$  that all closed curves have positive action and therefore  $\phi_{\eta,k}(x, x) \geq 0$  for all  $x \in M$ . But, if there existed  $x, y \in M$  such that  $\phi_{\eta,k}(x, y) = -\infty$ , then applying (1), we would get a contradiction:

$$\phi_{\eta,k}(x, x) \leq \phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) = -\infty.$$

Therefore,  $\phi_{\eta,k}(x, y) > -\infty$  for all  $x, y \in M$  if  $k \geq c(L)$ .

(3) Let  $k \geq c(L_\eta)$  and let  $Q := \max_{x \in M, \|v\|=1} L(x, v)$ . For any  $x, y \in M$  let us consider the unit speed geodesic connecting  $x$  to  $y$ ,  $\gamma_{x,y} : [0, d(x, y)] \longrightarrow M$ . Then:

$$\phi_{\eta,k}(x, y) \leq A_{L_\eta}(\gamma_{x,y}) \leq (Q + k)d(x, y).$$

Using (1) we can conclude:

$$\begin{aligned} \phi_{\eta,k}(x_2, y_2) - \phi_{\eta,k}(x_1, y_1) &= \phi_{\eta,k}(x_2, x_1) + \phi_{\eta,k}(x_1, y_1) + \phi_{\eta,k}(y_1, y_2) - \phi_{\eta,k}(x_1, y_1) \leq \\ &\leq (Q + k)[d(x_1, x_2) + d(y_1, y_2)]. \end{aligned}$$

(4) It follows immediately from (3).

(5) The first part is a consequence of (1) and (4). Let us prove the second part. Assume that  $k > c(L_\eta)$  and suppose by contradiction that there exist  $x \neq y \in M$

such that  $\phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) = 0$ . Consider  $\gamma_n \in C_{T_n}(x, y)$  and  $\sigma_n \in C_{S_n}(y, x)$  such that:

$$\lim_{n \rightarrow +\infty} A_{L_\eta}(\gamma_n) = \phi_{\eta,k}(x, y) \quad \text{and} \quad \lim_{n \rightarrow +\infty} A_{L_\eta}(\sigma_n) = \phi_{\eta,k}(y, x).$$

Let us prove that  $\underline{T} := \liminf_{n \rightarrow +\infty} T_n > 0$  (it might be  $+\infty$ ). Suppose by contradiction that  $\liminf_{n \rightarrow +\infty} T_n = 0$  and select a subsequence  $\{\gamma_{n_k}\}$  such that  $T_{n_k} \rightarrow 0$ . Using the superlinearity of  $L$ , we know that for each  $A > 0$  there exists  $B(A)$  such that  $L_\eta \geq A\|\dot{\gamma}\| - B(A)$ ; then:

$$\begin{aligned} \phi_{\eta,k}(x, y) &= \lim_{n_k \rightarrow \infty} \int_0^{T_{n_k}} L_\eta(\gamma_{n_k}, \dot{\gamma}_{n_k}) dt + kT_{n_k} \geq \\ &\geq \lim_{n_k \rightarrow \infty} \left[ A \int_0^{T_{n_k}} \|\dot{\gamma}_{n_k}\| dt + T_{n_k}(k - B(A)) \right] = \\ &= Ad(x, y). \end{aligned}$$

From the arbitrariness of  $A$  it follows that  $\phi_{\eta,k}(x, y) = +\infty$ , that is a contradiction. Therefore,  $\liminf_{n \rightarrow +\infty} T_n > 0$ .

Analogously one can prove that  $\underline{S} := \liminf_{n \rightarrow +\infty} S_n > 0$  (it might be  $+\infty$ ).

Choose now subsequences  $\{\gamma_{n_k}\}$  and  $\{\sigma_{m_k}\}$  such that  $T_{n_k} \rightarrow \underline{T} > 0$  and  $S_{m_k} \rightarrow \underline{S} > 0$ . Using the fact that  $k > c(L_\eta)$  and the fact that  $x \neq y \in M$  are such that  $\phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) = 0$ , we obtain a contradiction to (4) (or to the definition of  $c(L_\eta)$ ):

$$\begin{aligned} \phi_{\eta,c(L_\eta)}(x, x) &\leq \lim_{k \rightarrow +\infty} A_{L_\eta+c(L_\eta)}(\gamma_{n_k} * \sigma_{m_k}) \leq \\ &\leq \lim_{k \rightarrow +\infty} A_{L_\eta+k}(\gamma_{n_k} * \sigma_{m_k}) + \lim_{k \rightarrow +\infty} (c(L_\eta) - k)(T_{n_k} + S_{m_k}) = \\ &= \phi_{\eta,k}(x, y) + \phi_{\eta,k}(y, x) + \lim_{k \rightarrow +\infty} (c(L_\eta) - k)(T_{n_k} + S_{m_k}) = \\ &= (c(L_\eta) - k)(\underline{T} + \underline{S}) < 0 \quad (\text{or } -\infty). \end{aligned}$$

□

**Remark 5.10.** It follows from (2) that  $c(L_\eta)$  can be equivalently defined as:

$$\begin{aligned} c(L_\eta) &:= \inf\{k \in \mathbb{R} : \exists x, y \in M \text{ s.t. } \phi_{\eta,k}(x, y) > -\infty\} = \\ &= \sup\{k \in \mathbb{R} : \exists x, y \in M \text{ s.t. } \phi_{\eta,k}(x, y) = -\infty\}. \end{aligned}$$

In terms of Mañé potential, being a  $c$ -time free minimizer for  $L + k$  can be rewritten as:

$$\forall a < b \quad \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + (b - a)k = \phi_{\eta,k}(\gamma(a), \gamma(b)).$$

Of course, it does not make any sense to consider  $k < c(L_\eta)$ . Let us consider  $k \geq c(L_\eta)$ .

**Proposition 5.11.** *Let  $k > c(L_\eta)$ . For all  $x, y \in M$  with  $x \neq y$ , there exists  $T > 0$  and  $\gamma \in C_T(x, y)$  such that  $A_{L_\eta+k}(\gamma) = \phi_{\eta,k}(x, y)$ .*

**Proof.** Let define for  $T > 0$  the function  $f(T) := \min_{\gamma \in C_T(x, y)} A_{L_\eta+k}(\gamma)$ . This function is clearly continuous (for all  $k \geq c(L_\eta)$ ) and the following properties hold.

- $f(T) \rightarrow +\infty$  as  $T \rightarrow 0^+$  (this is true for all  $k \geq c(L_\eta)$ ). In fact, let  $\gamma_T$  be the corresponding Tonelli minimizer connecting  $x$  to  $y$  in time  $T$ . Using the superlinearity of  $L$ , for each  $A > 0$  there exists  $B = B(A)$  such that

$L(x, v) \geq A\|v\| - B$  for all  $(x, v)$ . Then:

$$\begin{aligned} f(T) &= \min_{\gamma \in C_T(x, y)} A_{L_\eta + k}(\gamma) = A_{L_\eta + k}(\gamma_T) \geq \\ &\geq A \int_0^T \|\dot{\gamma}_T\| dt + (k - B)T \geq \\ &= A d(x, y) + (k - B)T \xrightarrow{T \rightarrow 0^+} A d(x, y). \end{aligned}$$

Since  $A > 0$  is arbitrary and  $x \neq y$  then we can conclude that  $f(T) \xrightarrow{T \rightarrow 0^+} +\infty$ .

- $f(T) \rightarrow +\infty$  as  $t \rightarrow +\infty$  (this is true only for  $k > c(L_\eta)$ ). In fact:

$$\begin{aligned} f(T) &= \min_{\gamma \in C_T(x, y)} A_{L_\eta + k}(\gamma) = \\ &= \min_{\gamma \in C_T(x, y)} A_{L_\eta + c(L_\eta)}(\gamma) + (k - c(L_\eta))T \geq \\ &\geq \phi_{\eta, c(L_\eta)}(x, y) + (k - c(L_\eta))T \xrightarrow{T \rightarrow +\infty} +\infty. \end{aligned}$$

□

**Remark 5.12.** Hence if  $k > c(L_\eta)$ ,  $c$ -time free minimizers are not so special, since there are time free minimizers for  $L + k$  connecting any two given points, furthermore in a finite time. In the light of this (and other “*a-fortiori*” reasons), one should probably be more interested in studying the “critical” case  $k = c(L_\eta)$ , *i.e.*,  $c$ -time free minimizers for  $L + c(L_\eta)$ , that is for the least possible value of  $k$  for which they can exist.

**Definition 5.13 (c semi-static curves).** We say that  $\gamma : \mathbb{R} \rightarrow M$  is a  $c$  semi-static curve for  $L$  if:

$$\int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + c(L_\eta)(b - a) = \phi_{\eta, c(L_\eta)}(\gamma(a), \gamma(b)) \quad \forall a < b.$$

**Remark 5.14.** (i) If  $\gamma$  is  $c$  semi-static for  $L$ , then it is a  $c$ -time free minimizers for  $L + c(L_\eta)$  and consequently a  $c$ -global minimizer for  $L$ . Therefore, it corresponds to a solution of the Euler-Lagrange flow of  $L$ .

(ii) We shall prove in the following that for autonomous Tonelli Lagrangians, the converse is true: each  $c$ -global minimizer of  $L$  is indeed a  $c$  semi-static curve of  $L + c(L_\eta) = L + \alpha(c)$ , where  $\alpha$  denotes Mather’s  $\alpha$  function associated to  $L$  (see (13) in Section 4).

(iii) In Section 3 we proved that orbits on a KAM torus of cohomology class  $c$ , are  $c$ -time free minimizers for  $L + E_c$ , where  $E_c$  denoted the energy of the torus. It follows from Proposition 3.10 and the definition of Mañé critical value (see also Remark 5.10), that in this case  $E_c = c(L_c)$ . Therefore we can conclude that orbits on KAM tori are  $c$  semi-static. In particular, we can restate Remark 3.9 (iii) saying that if we have a KAM torus  $\mathcal{T}$  of cohomology class  $c$ , then

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is } c \text{ semi-static for } L \text{ and } t \in \mathbb{R}\},$$

where  $\mathcal{L}$  denotes the Legendre transform given by  $L$ .

Inspired by the last remark, we can define the following set.

**Definition 5.15 (Mañé set).** The Mañé set (with cohomology class  $c$ ) is:

$$\begin{aligned} \tilde{\mathcal{N}}_c &= \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a } c \text{ semi-static curve and } t \in \mathbb{R}\} = \\ (22) \quad &= \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a } c\text{-global minimizer and } t \in \mathbb{R}\}. \end{aligned}$$

**Remark 5.16.** The second equality in definition follows from Remark 5.14 and will be proved later on. Observe that so far we have not proved that such semi-static curves exist nor that this set is non-empty. We shall prove it later, deducing it - amongst other properties - from analogous results for another family of sets that we are about to define: the *Aubry sets*. However, if such set is non-empty, it is clearly invariant (it is union of orbits) and also closed (the proof follows the same line as Tonelli Theorem).

Let us start by recalling what happened in the case of orbits on a KAM torus. We saw in Section 3 that these orbits were not only  $c$ -global minimizers (or  $c$  semi-static), but they also satisfied a stronger property, stated in Proposition 3.11. Roughly speaking, the action of  $L_c + E_c$  on a piece of curve between two endpoints  $x$  and  $y$ , was not only the minimal needed to connect  $x$  to  $y$ , but it was also equal to *minus* the minimal action to connect  $y$  back to  $x$ . We called a curve satisfying such a condition a *regular minimizer*. Recall, in fact, that, as it follows easily from Proposition 5.9 (5), for each  $x, y \in M$  we have  $\phi_{\eta, c(L_\eta)}(x, y) \geq -\phi_{\eta, c(L_\eta)}(y, x)$ . Let us define such curves in the general case.

**Definition 5.17 (c static curves).** We say that  $\gamma : \mathbb{R} \rightarrow M$  is a  $c$  static curve for  $L$  (or a “ $c$ -regular minimizer”) if:

$$\int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + c(L_\eta)(b - a) = -\phi_{\eta, c(L_\eta)}(\gamma(b), \gamma(a)) \quad \forall a < b.$$

**Remark 5.18.** (i) If  $\gamma$  is  $c$  static for  $L$ , then it is a  $c$  semi-static (and therefore, as already observed before, it corresponds to a solution of the Euler-Lagrange flow of  $L$ ). It just a consequence of the fact that  $-\phi_{\eta, c(L_\eta)}(\gamma(b), \gamma(a)) \leq \phi_{\eta, c(L_\eta)}(\gamma(a), \gamma(b))$  and therefore if  $\gamma$  is  $c$ -static then:

$$\int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + c(L_\eta)(b - a) \leq \phi_{\eta, c(L_\eta)}(\gamma(a), \gamma(b)) \quad \forall a < b.$$

Since  $\phi_{\eta, c(L_\eta)}(\gamma(a), \gamma(b))$  was defined as the minimum over all connecting curves, then equality must hold.

(ii) Observe that the adjective *regular* in the alternative appellation (coined by John Mather) has no relation to the smoothness of the curve, since, like all solutions of the Euler-Lagrange flow, this curve will be as smooth as the Lagrangian.

(iii) In Section 3 (Proposition 3.11 and Remark 3.12) we proved that orbits on a KAM torus of cohomology class  $c$ , were  $c$ -static curves. Recall in fact that, as remarked before,  $E_c$ , which in that case denoted the energy of the torus, coincides with  $c(L_c)$  (use Proposition 3.10 and the definition of Mañé critical value or Remark 5.10). In particular, we can restate Remark 3.12 (ii) saying that if we have a KAM torus  $\mathcal{T}$  of cohomology class  $c$ , then

$$\mathcal{L}^{-1}(\mathcal{T}) = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is } c\text{-static for } L \text{ and } t \in \mathbb{R}\},$$

where  $\mathcal{L}$  denotes the Legendre transform given by  $L$ .

(iv) The names *semi-static* and *static* are probably inspired by the fact that in the case of mechanical Lagrangians (see Section 2), 0-static curves correspond to “some” fixed points of the flow (namely the minima of the potential), while 0-semi static curves may possibly include also hetero/homoclinic connections amongst these fixed points (see Addendum 5B on the pendulum).

Inspired by the last remark, we define the following set.

**Definition 5.19 (Aubry set).** The Aubry set (with cohomology class  $c$ ) is:

$$(23) \quad \tilde{\mathcal{A}}_c = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a } c \text{ static curve and } t \in \mathbb{R}\}.$$

The projection on the base manifold  $\mathcal{A}_c = \pi(\tilde{\mathcal{A}}_c) \subseteq M$  is called *projected Aubry set* (with cohomology class  $c$ ).

**Remark 5.20.** Observe that so far we have not proved that this set is non-empty! We shall do it now. However, if such set is non-empty, it is clearly invariant (it is union of orbits) and also closed (the proof follows the same line as Tonelli Theorem). Moreover,  $\tilde{\mathcal{A}}_c$  is clearly contained in  $\tilde{\mathcal{N}}_c$  (Remark 5.18 (i)).

We summarise in this diagram what we are going to prove in the remaining of this section.

**Theorem 5.21.** <sup>1</sup> (1), (2), (3), (4), (5) and (6) in the following diagram are true.

$$\begin{array}{ccccccc}
 \tilde{\mathcal{M}}_c & \stackrel{(1)}{\subseteq} & \tilde{\mathcal{A}}_c & \stackrel{(2)}{\subseteq} & \tilde{\mathcal{N}}_c & \stackrel{(3)}{\subseteq} & \tilde{\mathcal{E}}_c := \{E(x, v) = \alpha(c) \stackrel{(4)}{=} c(L_\eta)\} \subseteq TM \\
 \uparrow \downarrow & & \uparrow \downarrow & & & & \downarrow \pi \\
 (\pi|_{\tilde{\mathcal{M}}_c})^{-1} & \stackrel{(5)}{\subseteq} & \pi & \stackrel{(6)}{\subseteq} & (\pi|_{\tilde{\mathcal{A}}_c})^{-1} & & \\
 \mathcal{M}_c & \subseteq & \mathcal{A}_c & \subseteq & & \subseteq & M
 \end{array}$$

**Remark 5.22.** (i) It follows from (1), (2) and the existence of  $c$ -action minimizing measures (Corollary 4.5) that the Aubry and Mañé sets are non-empty. Therefore, there exist semi-static and static curves.

(ii) Inclusion (2), as already observed, follows obviously from the fact that static curves are also semi-static (Remark 5.18 (i)).

(iii) The above inclusions (1) and (2) may not be strict (see Remark 5.42 in Addendum 5B).

(iv) Since  $\tilde{\mathcal{A}}_c$  and  $\tilde{\mathcal{N}}_c$  are closed, it follows from (3) that they are compact.

(v) The proof of (3) provides a proof of Carneiro's theorem stated in Section 4 (Theorem 4.13).

(vi) Properties (5) and (6) are what are generally called Mather's graph theorem(s). Namely, the Mather set and the Aubry set are contained in a Lipschitz graph over  $M$ . This is probably the most important property of these sets and it has many dynamical consequences. In some sense, this is why they can be thought as generalization of KAM tori (or Lagrangian graphs).

(vii) The graph property does not hold in general for the Mañé set (see Addendum 5B on the pendulum).

Let us start by proving that  $c$  semi-static curves have energy  $c(L_\eta)$ . As we already remarked above, the original version of this theorem (with  $c(L_\eta)$  replaced by  $\alpha(c)$ ) is due to Carneio (see Theorem 4.13). The proof presented here follows an idea of Ricardo Mañé [35] (see also [14, Theorem XI])

**Proposition 5.23** (Property (3)).  $\tilde{\mathcal{N}}_c \subseteq \tilde{\mathcal{E}}_c := \{E(x, v) = c(L_\eta)\}$ , i.e.,  $c$  semi-static curves have energy equal to  $c(L_\eta)$ .

**Proof.** Let  $\gamma : \mathbb{R} \rightarrow M$  be a  $c$  semi-static curve, i.e., for each  $T > 0$  we have  $A_{L_\eta + c(L_\eta)}(\gamma|_{[0, T]}) = \phi_{\eta, c(L_\eta)}(\gamma(0), \gamma(T))$ .

Let us fix  $T > 0$  and  $\lambda > 0$ , and consider a time-reparameterization of  $\gamma$ , given by  $\gamma_\lambda : [0, T/\lambda] \rightarrow M$ ,  $t \mapsto \gamma(\lambda t)$ . Observe that the end-points are not changed:  $\gamma_\lambda(0) = \gamma(0)$  and  $\gamma_\lambda(T/\lambda) = \gamma(T)$ , but only the time length. Let us consider the

<sup>1</sup>This (unintentional?) typographical "coincidence" honoring Ricardo Mañé was first pointed out by Albert Fathi.

action of these curves as a function of  $\lambda > 0$ :

$$\begin{aligned} A(\lambda) &:= A_{L_\eta + c(L_\eta)}(\gamma_\lambda) = \int_0^{T/\lambda} L_\eta(\gamma_\lambda(t), \dot{\gamma}_\lambda(t)) dt + c(L_\eta)T/\lambda = \\ &= \int_0^{Tl} L_\eta(\gamma(\lambda t), \lambda \dot{\gamma}(\lambda t)) dt + c(L_\eta)T/\lambda. \end{aligned}$$

Since  $\gamma$  is  $c$  semi-static (and therefore it is a  $c$ -time free minimizer for  $L_\eta + c(L_\eta)$ ), then  $A$  has a minimum at  $\lambda = 1$ . Therefore:

$$\begin{aligned} 0 = A'(1) &= -T[L_\eta(\gamma(T), \dot{\gamma}(T)) + c(L_\eta)T] + \\ &+ \int_0^T \left( \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t))\dot{\gamma}(t) + \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t))[\dot{\gamma}(t) + t\ddot{\gamma}(t)] \right) dt. \end{aligned}$$

Integrating by parts (observe that  $\frac{d}{dt}L = (\frac{\partial L}{\partial x}\dot{\gamma} + \frac{\partial L}{\partial v}\ddot{\gamma})t$ ) and recalling the definition of the energy  $E(x, v) = \frac{\partial L}{\partial v}(x, v)v - L(x, v)$  and the fact that is preserved along the orbit ( $E(\gamma(t), \dot{\gamma}(t)) = E(\gamma(0), \dot{\gamma}(0))$  for all  $t \in \mathbb{R}$ ), we obtain:

$$\begin{aligned} 0 &= A'(1) = \dots = \\ &= -T[L_\eta(\gamma(T), \dot{\gamma}(T)) + c(L_\eta)] + L_\eta(\gamma(t), \dot{\gamma}(t)) \Big|_0^T - \\ &\quad - \int_0^T \left( -L(\gamma(t), \dot{\gamma}(t)) + \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t))\dot{\gamma}(t) \right) dt = \\ &= -Tc(L_\eta) + \int_0^T E(\gamma(t), \dot{\gamma}(t)) dt = \\ &= [E(\gamma(0), \dot{\gamma}(0)) - c(L_\eta)]T. \end{aligned}$$

Hence,  $E(\gamma(0), \dot{\gamma}(0)) = c(L_\eta)$ .  $\square$

Let us now prove that Mañé critical value coincides with Mather's  $\alpha$ -function. We shall follow the proof given in [14, Theorem II]

**Proposition 5.24** (Property (4)).  $c(L_\eta) = \alpha(c)$ , where  $c = [\eta]$ .

**Proof.** Suppose that  $\mu$  is an invariant ergodic probability measure. If we fix a generic point  $(x, v)$  in the support of  $\mu$ , it follows from the ergodic theorem that there exists a sequence of times  $T_n \rightarrow +\infty$  such that  $\Phi_{T_n}^L(x, v) \rightarrow (x, v)$  as  $T_n \rightarrow +\infty$  and

$$\int L_\eta d\mu = \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) dt.$$

For the sake of simplifying the notation, let us denote  $(x_t, v_t) := \Phi_t^L(x, v)$ . Let  $B := \max\{|L_\eta(x, v)| : \|v\| \leq 1\}$  and for each  $n$  denote by  $\sigma_n : [0, d(x, x_{T_n})] \rightarrow M$  the geodesic joining  $x$  to  $x_{T_n}$  and by  $\gamma_n : [0, T_n] \rightarrow M$  the projection of the orbit (*i.e.*,  $\gamma_n(t) = x_t$ ). We have:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{T_n} A_{L_\eta + k}(\gamma_n * \sigma_n) &= \\ &= \lim_{n \rightarrow \infty} \left( \frac{1}{T_n} A_{L_\eta + k}(\gamma_n) + \frac{1}{T_n} \int_0^{d(x, x_{T_n})} L_\eta(\sigma_n(t), \dot{\sigma}_n(t)) + k dt \right) \leq \\ &= \leq \lim_{n \rightarrow \infty} \left( \frac{1}{T_n} A_{L_\eta + k}(\gamma_n) + \frac{1}{T_n} (B + k) \text{diam}(M) \right) \leq \\ &= A_{L_\eta}(\mu) + k. \end{aligned}$$



Therefore, if  $k < -A_{L_\eta}(\mu)$  then  $\Phi_k(x, x) = -\infty$ . Hence,  $k \leq c(L_\eta)$ . It follows that:

$$\begin{aligned} c(L_\eta) &\geq \sup\{-A_{L_\eta}(\mu), \mu \in \mathfrak{M}_{\text{erg}}(L)\} \geq \\ &\geq -\inf\{A_{L_\eta}(\mu), \mu \in \mathfrak{M}(L)\} = \alpha(c). \end{aligned}$$

Now we want to prove the reversed inequality. Let  $k < c(L_\eta)$  and  $x, y \in M$ . Since  $\Phi_k(x, y) = -\infty$ , there exists a sequence of absolutely continuous curves  $\gamma_n : [0, T_n] \rightarrow M$  such that  $\gamma_n(0) = x$ ,  $\gamma_n(T_n) = y$  (i.e.,  $\gamma_n \in C_{T_n}(x, y)$ ) and

$$\lim_{n \rightarrow \infty} A_{L_\eta+k}(\gamma_n) = -\infty.$$

Moreover, since  $L$  is bounded from below, we have that  $T_n \rightarrow +\infty$ . Let now  $y_n$  be a Tonelli minimizer in  $C_{T_n}(x, y)$ . Then, consider the invariant probability measure evenly distributed along  $\gamma_n$ . The family of these measures is pre-compact and we can extract a subsequence converging (in the weak\* topology) to an invariant probability measure  $\mu$ . In particular,

$$A_{L_\eta}(\mu) + k = \lim_{n \rightarrow \infty} \frac{1}{T_n} A_{L_\eta+k}(\gamma_n).$$

But since  $T_n > 0$  and  $\lim_{n \rightarrow \infty} A_{L_\eta+k}(\gamma_n) = \Phi_k(x, y) = -\infty$ , then  $A_{L_\eta}(\mu) + k \leq 0$ . Therefore, for any  $k < c(L_\eta)$  we can find an invariant probability measure  $\mu$  such that  $k \leq -A_{L_\eta}(\mu)$ . Therefore:

$$\begin{aligned} c(L_\eta) &\leq \sup\{-A_{L_\eta}(\mu), \mu \in \mathfrak{M}_{\text{erg}}(L)\} \geq \\ &\leq -\inf\{A_{L_\eta}(\mu), \mu \in \mathfrak{M}(L)\} = \alpha(c). \end{aligned}$$

□

We shall provide two other alternative proofs of the above proposition. The first one uses the fact that the Mather set is included into the Mañé set (but observe that in our proof of this fact - Proposition 5.26 - we use that  $\alpha(c) = c(L_\eta)$ ). The second one is shorter but it requires the use of holonomic measures (see Addendum 4C). See also [14, Theorem II] for another proof.

**Alternative Proof I.** Since  $\tilde{\mathcal{N}}_c$  is compact (it follows from (3)) and invariant under the Euler-Lagrange flow  $\Phi_t^L$ , then there exists an invariant ergodic probability measure  $\mu$  supported in it (Kryloff and Bogoliouboff [31], compare also with Proposition 4.1). If we fix a generic point  $(x, v)$  in the support of  $\mu$ , it follows from the ergodic theorem that there exists a sequence of times  $T_n \rightarrow +\infty$  such that  $\Phi_{T_n}^L(x, v) \rightarrow (x, v)$  as  $T_n \rightarrow +\infty$  and

$$\int L_\eta d\mu = \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) dt.$$

Then, using the definition of  $\alpha(c)$  and the fact that orbits in the support of this measure are semi-static, we obtain:

$$\begin{aligned} -\alpha(c) &\leq \int L_\eta d\mu = \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) dt = \\ &= \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) + c(L_\eta) dt - c(L_\eta) = \\ &= \lim_{n \rightarrow +\infty} \frac{\phi_{\eta, c(L_\eta)}(x, \pi(\Phi_{T_n}^L(x, v)))}{T_n} - c(L_\eta) = -c(L_\eta), \end{aligned}$$

where in the last equality we used that  $\phi_{\eta, c(L_\eta)}$  is bounded on  $M \times M$  (being Lipschitz on a compact manifold). Therefore  $\alpha(c) \geq c(L_\eta)$ .

For the reversed inequality, we shall use the following version of the Ergodic theorem (see for instance [34, Lemma 2.1] for a proof).

**Lemma 5.25.** *Let  $(X, d)$  be a complete metric space and  $(X, \mathcal{B}, \nu)$  a probability space. Let  $f$  be an ergodic measure preserving map and  $F : X \rightarrow \mathbb{R}$  a  $\nu$ -integrable function. Then, for  $\nu$ -almost every  $x \in X$  the following property holds:*

$$\forall \varepsilon > 0 \quad \exists N > 0 : \quad d(f^N(x), x) < \varepsilon \quad \text{and} \quad \left| \sum_{j=0}^{N-1} F(f^j(x)) - N \int F d\nu \right| < \varepsilon.$$

Let us now see how to use this lemma for our purposes. Let  $\mu$  be a  $c$ -action minimizing ergodic measure, *i.e.*,

$$\int (L_\eta(x, v) + \alpha(c)) d\mu = 0.$$

Applying the above Lemma with  $F = L_\eta + \alpha(c)$  and  $X = TM$ , we obtain that there exists a  $\mu$ -full measure set  $A$  such that if  $(x, v) \in A$ , then there exists a sequence  $T_n \rightarrow +\infty$  such that

$$d((x, v), \Phi_{T_n}^L(x, v)) \xrightarrow{n \rightarrow +\infty} 0 \quad \text{and} \quad \int_0^{T_n} (L_\eta(\Phi_t^L(x, v)) + \alpha(c)) dt \xrightarrow{n \rightarrow +\infty} 0.$$

Then:

$$\begin{aligned} & \phi_{\eta, \alpha(c)}(x, \pi(\Phi_1^L(x, v))) + \phi_{\eta, \alpha(c)}(\pi(\Phi_1^L(x, v)), x) = \\ &= \lim_{n \rightarrow \infty} (\phi_{\eta, \alpha(c)}(x, \pi(\Phi_1^L(x, v))) + \phi_{\eta, \alpha(c)}(\pi(\Phi_1^L(x, v)), \pi(\Phi_{T_n}^L(x, v)))) \leq \\ &\leq \lim_{n \rightarrow \infty} \int_0^{T_n} (L_\eta(\Phi_t^L(x, v)) + \alpha(c)) dt = 0. \end{aligned}$$

It follows from the second property in Proposition 5.9 (5) that  $\alpha(c) \leq c(L_\eta)$  and this concludes the proof.  $\square$

**Alternative Proof II.** Let  $\gamma$  be any closed curve and let  $\mu_\gamma$  be the probability measure evenly distributed on it (see Addendum 4C). Let  $k \geq c(L_\eta)$ . It follows from the definition of  $c(L_\eta)$  that  $A_{L_\eta+k}(\mu_\gamma) \geq 0$  and therefore  $A_{L_\eta}(\mu_\gamma) \geq -k$ . It follows from the definition of holonomic measure then, that for any  $\mu$  holonomic probability measure, we have  $A_{L_\eta}(\mu) \geq -k$ . Taking the infimum over all holonomic measure and using the result mentioned in iv) in Addendum 4C, we can conclude that  $-\alpha(c) \geq -k$ . Since this holds for all  $k \geq c(L_\eta)$ , we obtain:  $\alpha(c) \leq c(L_\eta)$ .

To prove the reversed inequality, observe that if  $k < c(L_\eta)$ , then there exists  $\gamma$  closed curve such that  $A_{L_\eta+k}(\gamma) < 0$ . Therefore, if  $\mu_\gamma$  is the associated holonomic measure we obtain:

$$-\alpha(c) \leq A_{L_\eta}(\mu_\gamma) < -k.$$

Since this holds for all  $k < c(L_\eta)$ , we conclude that  $\alpha(c) \geq c(L_\eta)$ .  $\square$

The proof of Property (1) is essentially similar to the (first) proof of Proposition 5.24. We shall prove this more general result, due to Ricardo Mañé [35] (see also [14, Theorem IV]).

**Proposition 5.26** (Property (1)).  *$\mu \in \mathfrak{M}(L)$  is  $c$ -action minimizing if and only if  $\text{supp } \mu \subseteq \tilde{\mathcal{A}}_c$ . In particular,  $\tilde{\mathcal{M}}_c \subseteq \tilde{\mathcal{A}}_c$ .*

**Proof.** Since  $\tilde{\mathcal{A}}_c$  is closed, it is sufficient to prove the results only for the ergodic measures.

[ $\Leftarrow$ ] Let  $\mu \in \mathfrak{M}(L)$  be ergodic and suppose that  $\text{supp } \mu \subseteq \tilde{\mathcal{A}}_c$ . Applying the ergodic theorem, we know that for a  $\mu$ -generic point  $(x, v)$  in the support of  $\mu$ , there

exists a sequence of times  $T_n \rightarrow +\infty$  such that  $\Phi_{T_n}^L(x, v) \rightarrow (x, v)$  as  $T_n \rightarrow +\infty$  and

$$\int L_\eta d\mu = \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) dt.$$

Then, using the fact that orbits in the support of this measure are semi-static and that  $\alpha(c) = c(L_\eta)$ , we obtain:

$$\begin{aligned} \int L_\eta + \alpha(c) d\mu &= \int L_\eta + c(L_\eta) d\mu = \\ &= \lim_{n \rightarrow +\infty} \frac{1}{T_n} \int_0^{T_n} L_\eta(\Phi_t^L(x, v)) + c(L_\eta) dt = \\ &= \lim_{n \rightarrow +\infty} \frac{-\phi_{\eta, c(L_\eta)}(\pi(\Phi_{T_n}^L(x, v)), x)}{T_n} = 0, \end{aligned}$$

where in the last equality we used that  $\phi_{\eta, c(L_\eta)}$  is bounded on  $M \times M$  (being Lipschitz on a compact manifold). Therefore  $\int L_\eta d\mu \leq -\alpha(c)$  and from the definition of  $\alpha(c)$ , it follows that  $\int L_\eta d\mu = -\alpha(c)$ , i.e.,  $\mu$  is  $c$ -action minimizing.

[ $\Rightarrow$ ] We shall use the above mentioned version of the ergodic theorem (see Lemma 5.25).

Let  $\mu$  be a  $c$ -action minimizing ergodic measure, i.e.,

$$\int (L_\eta(x, v) + \alpha(c)) d\mu = 0.$$

Applying the above Lemma with  $F = L_\eta + \alpha(c)$  and  $X = TM$ , we obtain that there exists a  $\mu$ -full measure set  $A$  such that if  $(x, v) \in A$ , then there exists a sequence  $T_n \rightarrow +\infty$  such that

$$d((x, v), \Phi_{T_n}^L(x, v)) \xrightarrow{n \rightarrow +\infty} 0 \quad \text{and} \quad \int_0^{T_n} (L(\Phi_t^L(x, v)) + \alpha(c)) dt \xrightarrow{n \rightarrow +\infty} 0.$$

Then, let  $a > 0$  and recall that  $\alpha(c) = c(L_\eta)$ :

$$\begin{aligned} &\phi_{\eta, c(L_\eta)}(x, \pi(\Phi_a^L(x, v))) + \phi_{\eta, c(L_\eta)}(\pi(\Phi_a^L(x, v)), x) = \\ &= \lim_{n \rightarrow \infty} (\phi_{\eta, c(L_\eta)}(x, \pi(\Phi_a^L(x, v))) + \phi_{\eta, c(L_\eta)}(\pi(\Phi_a^L(x, v)), \pi(\Phi_{T_n}^L(x, v)))) \leq \\ &\leq \lim_{n \rightarrow \infty} \int_0^{T_n} (L_\eta(\Phi_t^L(x, v)) + c(L_\eta)) dt = \\ &\leq \lim_{n \rightarrow \infty} \int_0^{T_n} (L_\eta(\Phi_t^L(x, v)) + \alpha(c)) dt = 0. \end{aligned}$$

Recalling property (5) in Proposition 5.9, we can conclude that for any  $a > 0$ :

$$\phi_{\eta, c(L_\eta)}(x, \pi(\Phi_a^L(x, v))) + \phi_{\eta, c(L_\eta)}(\pi(\Phi_a^L(x, v)), x) = 0$$

and therefore the orbit through the point  $(x, v)$  is  $c$ -static. Since the points for which this reasoning can be applied are dense in the support of  $\mu$  and  $\tilde{\mathcal{A}}_c$  is closed, then we prove the claim:  $\text{supp } \mu \subseteq \tilde{\mathcal{A}}_c$ .  $\square$

**Remark 5.27.** Looking at the proof of Proposition 5.26, it is quite easy to see that we actually proved that: if  $\mu \in \mathfrak{M}(L)$  is such that  $\text{supp } \mu \subseteq \tilde{\mathcal{N}}_c$ , then  $\mu$  is  $c$ -action minimizing (check that the same proof still works in this case). In fact, one can prove this stronger version:

**Proposition 5.28.**  $\mu \in \mathfrak{M}(L)$  is  $c$ -action minimizing if and only if  $\text{supp } \mu \subseteq \tilde{\mathcal{N}}_c$ .

We shall deduce such a proposition from the fact (to be proved later in Section 6, Proposition 6.33) that the “non-wandering set” of the Mañé’ set is contained in the Aubry set (see also [14, Theorem V.c]).

It is also quite easy to check that orbits in the Mañé (resp. Aubry) set are asymptotic to the Mather set.

**Proposition 5.29.** *If  $\gamma : \mathbb{R} \rightarrow M$  is a  $c$  semi-static curve, then*

$$\liminf_{t \rightarrow \pm\infty} d(\gamma(t), \mathcal{M}_c) = 0.$$

**Proof.** Let  $T \geq 1$  and consider the probability measure  $\mu_T$  evenly distributed along the piece of curve  $\{(\gamma(t), \dot{\gamma}(t)) : t \in [0, T]\}$  (for a definition, for instance, the reader may check the proof of Proposition 4.1). The Lagrangian actions of these measures are equi-bounded (we use here that the orbit is semi-static):

$$\begin{aligned} A_{L_\eta}(\mu_T) &= \frac{1}{T} \int_0^T L_\eta(\gamma(t), \dot{\gamma}(t)) dt = \frac{\phi_{\eta, c(L_\eta)}(\gamma(0), \gamma(T))}{T} - c(L_\eta) \leq \\ &\leq \max_{M \times M} \phi_{\eta, c(L_\eta)}(x, y) - c(L_\eta) < \infty. \end{aligned}$$

Therefore, this family of measures is pre-compact with respect to the weak\* topology. Let us consider any converging subsequence  $\mu_{T_k} \rightarrow \mu$ , with  $T_k \rightarrow +\infty$ . Then,  $\mu$  is invariant (see, again, the proof of Proposition 4.1) and:

$$\begin{aligned} \int L_\eta d\mu &= \lim_{k \rightarrow \infty} \int L_\eta d\mu_k = \lim_{k \rightarrow \infty} \frac{\phi_{\eta, c(L_\eta)}(\gamma(0), \gamma(T_k))}{T_k} - c(L_\eta) = \\ &= -c(L_\eta) = -\alpha(c). \end{aligned}$$

Therefore,  $\mu$  is  $c$ -action minimizing.  $\square$

Finally, we prove the most important result of this theory: the graph property of the Mather and Aubry sets, respectively, Property (5) and (6) in the diagram (see also Theorem 4.12). Because of the inclusion proved in Proposition 5.26, it is sufficient to prove it for the Aubry set.

**Theorem 5.30 (Mather's graph theorem, [42]).** (Property (6))  $\pi|_{\tilde{\mathcal{A}}_c}$  is an injective mapping of  $\tilde{\mathcal{A}}_c$  into  $M$ , and its inverse  $(\pi|_{\tilde{\mathcal{A}}_c})^{-1} : \mathcal{A}_c \rightarrow \tilde{\mathcal{A}}_c$  is Lipschitz.

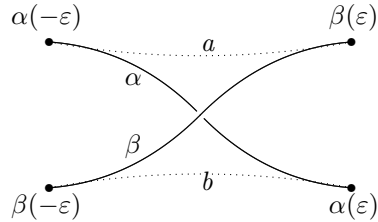
The proof of this theorem will be based on the following “crossing” Lemma, proved by Mather in [42] (to which we refer the reader for a complete proof).

**Lemma 5.31 (Mather's crossing lemma, [42]).** *Let  $K > 0$ . There exist  $\varepsilon, \delta, \vartheta > 0$ ,  $C > 0$  such that if  $\alpha, \beta : [-\varepsilon, \varepsilon] \rightarrow M$  are solutions of the Euler-Lagrange equation with  $\|(\alpha(0), \dot{\alpha}(0))\| \leq K$ ,  $\|(\beta(0), \dot{\beta}(0))\| \leq K$  and*

$$d(\alpha(0), \beta(0)) \leq \delta \quad \text{and} \quad d\left((\alpha(0), \dot{\alpha}(0)), (\beta(0), \dot{\beta}(0))\right) > Cd(\alpha(0), \beta(0)),$$

*then there exist  $C^1$  curves  $a, b : [-\varepsilon, \varepsilon] \rightarrow M$  with end-points  $a(-\varepsilon) = \alpha(-\varepsilon)$ ,  $a(\varepsilon) = \beta(\varepsilon)$  and  $b(-\varepsilon) = \beta(-\varepsilon)$ ,  $b(\varepsilon) = \alpha(\varepsilon)$  such that:*

$$A_{L_\eta}(\alpha) + A_{L_\eta}(\beta) - A_{L_\eta}(a) - A_{L_\eta}(b) \geq \vartheta d\left((\alpha(0), \dot{\alpha}(0)), (\beta(0), \dot{\beta}(0))\right)^2 > 0.$$



**Proof.** [Theorem 5.30] We shall first use Lemma 5.31 to prove the Lipschitz property and then the rest will be just a consequence of this property. Let choose  $K := \max_{\tilde{\mathcal{A}}_c} \|(x, v)\|$  (this is finite since  $\tilde{\mathcal{A}}_c$  is compact) and let  $\varepsilon, \delta, \vartheta, C$  be as in Lemma 5.31. Then, we shall prove that: *if  $(x_1, v_1), (x_2, v_2) \in \tilde{\mathcal{A}}_c$  are such that  $d(x_1, x_2) \leq \delta$ , then  $d((x_1, v_1), (x_2, v_2)) \leq Cd(x_1, x_2)$ .*

Suppose by contradiction that  $d((x_1, v_1), (x_2, v_2)) > Cd(x_1, x_2)$  and consider the flow lines through these points, namely  $\alpha(t) := \Phi_t^L(x_1, v_1)$  and  $\beta(t) := \Phi_t^L(x_2, v_2)$ . They satisfy the hypothesis of Lemma 5.31 (with our choice of  $K$ ) and hence we can deduce the existence of two other curves  $a, b : [-\varepsilon, \varepsilon] \rightarrow M$  with end-points  $a(-\varepsilon) = \alpha(-\varepsilon)$ ,  $a(\varepsilon) = \beta(\varepsilon)$  and  $b(-\varepsilon) = \beta(-\varepsilon)$ ,  $b(\varepsilon) = \alpha(\varepsilon)$ , such that:

$$A_{L_\eta}(a) + A_{L_\eta}(b) < A_{L_\eta}(\alpha) + A_{L_\eta}(\beta).$$

But then:

$$\begin{aligned} & \phi_{\eta, c(L_\eta)}(a(-\varepsilon), a(\varepsilon)) + \phi_{\eta, c(L_\eta)}(b(-\varepsilon), b(\varepsilon)) \leq \\ & \leq A_{L_\eta}(a) + A_{L_\eta}(b) < A_{L_\eta}(\alpha) + A_{L_\eta}(\beta) = \\ & = -\phi_{\eta, c(L_\eta)}(\alpha(\varepsilon), \alpha(-\varepsilon)) - \phi_{\eta, c(L_\eta)}(\beta(\varepsilon), \beta(-\varepsilon)). \end{aligned}$$

where in the last equality we used that  $\alpha$  and  $\beta$  are  $c$ -semistatic. The above inequality and the triangle inequality for  $\phi_{\eta, c(L_\eta)}$  (see Proposition 5.9 (1)) lead to a contradiction to Proposition 5.9 (5):

$$\begin{aligned} & \phi_{\eta, c(L_\eta)}(\alpha(-\varepsilon), \beta(\varepsilon)) = \phi_{\eta, c(L_\eta)}(a(-\varepsilon), a(\varepsilon)) < \\ & < -(\phi_{\eta, c(L_\eta)}(b(-\varepsilon), b(\varepsilon)) + \phi_{\eta, c(L_\eta)}(\alpha(\varepsilon), \alpha(-\varepsilon)) + \phi_{\eta, c(L_\eta)}(\beta(\varepsilon), \beta(-\varepsilon))) = \\ & = -(\phi_{\eta, c(L_\eta)}(\beta(-\varepsilon), \alpha(\varepsilon)) + \phi_{\eta, c(L_\eta)}(\alpha(\varepsilon), \alpha(-\varepsilon)) + \phi_{\eta, c(L_\eta)}(\beta(\varepsilon), \beta(-\varepsilon))) \leq \\ & \leq -(\phi_{\eta, c(L_\eta)}(\beta(-\varepsilon), \alpha(-\varepsilon)) + \phi_{\eta, c(L_\eta)}(\beta(\varepsilon), \beta(-\varepsilon))) \leq \\ & \leq -\phi_{\eta, c(L_\eta)}(\beta(\varepsilon), \alpha(-\varepsilon)). \end{aligned}$$

Therefore the inverse of the projection is locally Lipschitz and this concludes the proof.  $\square$

**Remark 5.32.** (i) Actually it follows from the proof (in the choice of  $K$ ) that the graphs of the Aubry sets (or Mather sets) corresponding to compact sets of cohomology classes, are equi-Lipschitz.

(ii) An alternative proof of the graph property will be presented in Section 6, following Fathi's weak KAM theory.

One can show several other properties of these sets. For instance, we have remarked that the Mather sets, being the support of invariant probability measures, are recurrent under the flow. This is not true anymore for the Aubry and Mañé sets, but something can still be said. Let us first recall the definition of  $\varepsilon$ -pseudo orbit. Given a (compact) metric space  $X$  and a flow  $\varphi$  on it, we say that there exists an  $\varepsilon$ -pseudo orbit between two points  $x, y \in X$ , if we can find  $\{x_n\}_{n=0}^{k_\varepsilon} \subset X$  and positive times  $t_1, \dots, t_{k_\varepsilon} > 0$  such that  $x_0 = x$ ,  $x_{k_\varepsilon} = y$  and  $\text{dist}(\varphi_{t_{i+1}}(x_i), x_{i+1}) \leq \varepsilon$  for all  $i = 0, \dots, k_\varepsilon$ .

**Proposition 5.33.** (i)  $\Phi^L|_{\tilde{\mathcal{N}}_c}$  is chain transitive, i.e., for each  $\varepsilon > 0$  and for all  $(x, v), (y, w) \in \tilde{\mathcal{N}}_c$ , there exists an  $\varepsilon$ -pseudo-orbit for the flow  $\Phi^L$  connecting them. (ii)  $\Phi^L|_{\tilde{\mathcal{A}}_c}$  is chain recurrent, i.e., for each  $\varepsilon > 0$  and for all  $(x, v) \in \tilde{\mathcal{A}}_c$ , there exists an  $\varepsilon$ -pseudo-orbit for the flow  $\Phi^L$  connecting  $(x, v)$  to itself.

The proof of this result can be found for instance in [14, Theorem V]

As a consequence of the chain-transitivity it follows that the Mañé set must be connected (the Aubry set in general not).

**Corollary 5.34.** *The Mañé set is connected*

## ADDENDA

### 5.A - SOME TOPOLOGICAL AND SYMPLECTIC PROPERTIES OF THESE SETS

In this Addendum we want to discuss (without any proof) some topological and symplectic properties of the Aubry and Mañé sets, similar to what we have already seen and proved for the Mather sets.

In Section 5, Proposition 4.27, we had related the intersection of Mather sets corresponding to different cohomology classes, to the “flatness” of the  $\alpha$  function. The same result holds for the Aubry set and has been proved by Daniel Massart in [37, Proposition 6]. However, the proof in this case is less straightforward and more involved.

**Proposition 5.35 (Massart, [37]).** *Let  $c \in H^1(M; \mathbb{R})$  and denote by  $F_c$  maximal face of the epigraph of  $\alpha$  containing  $c$  in its interior.*

- (i) *If a cohomology class  $c_1$  belongs  $F_c$ , then  $\mathcal{A}_c \subseteq \mathcal{A}_{c_1}$ . In particular, if  $c_1$  belongs to the interior of  $F_c$ , then they coincide, i.e.,  $\mathcal{A}_c = \mathcal{A}_{c_1}$ .*
- (ii) *Conversely, if two cohomology classes  $c$  and  $c_1$  are such that  $\tilde{\mathcal{A}}_c \cap \tilde{\mathcal{A}}_{c_1} \neq \emptyset$ , then for each  $\lambda \in [0, 1]$  we have  $\alpha(c) = \alpha(\lambda c + (1 - \lambda)c_1)$ , i.e., the epigraph of  $\alpha$  has a face containing  $c$  and  $c_1$ .*

In particular, Massart proved that it is possible to relate the dimension of a “face” of the epigraph of the  $\alpha$ -function to the topological complexity of the Aubry sets corresponding to cohomologies in that face (see [37, Theorem 1]). More precisely, for any sufficiently small  $\varepsilon > 0$ , let us define  $C_c(\varepsilon)$  be the set of integer homology classes which are represented by a piecewise  $C^1$  closed curve made with arcs contained in  $\mathcal{A}_c$  except for a remainder of total length less than  $\varepsilon$ . Let  $C_c := \bigcap_{\varepsilon > 0} C_c(\varepsilon)$ . Let  $V_c$  be the space spanned in  $H_1(M; \mathbb{R})$  by  $C_c$ . Note that  $V_c$  is an integer subspace of  $H_1(M; \mathbb{R})$ , that is it has a basis of integer elements (images in  $H_1(M; \mathbb{R})$  of elements in  $H_1(M; \mathbb{Z})$ ).

We denote by:

- $F_c$  the maximal face (flat piece) of the epigraph of  $\alpha$ , containing  $c$  in its interior;
- $\text{Vect } F_c$  the underlying vector space of the affine subspace generated by  $F_c$  in  $H^1(M; \mathbb{R})$ ;
- $V_c^\perp$  the vector space of cohomology classes of  $C^1$  1-forms that vanish on  $V_c$ ;
- $G_c$  the vector space of cohomology classes of  $C^1$  1-forms that vanish in  $T_x M$  for each  $x \in \mathcal{A}_c$ ;
- $E_c$  the space of cohomology classes of 1-forms of class  $C^1$ , the supports of which are disjoint from  $\mathcal{A}_c$ .

**Theorem 5.36 (Massart, [37]).**  $E_c \subseteq \text{Vect } F_c \subseteq G_c \subseteq V_c^\perp$ .

Moreover, as we have already proved for the Mather sets (see Addendum 4A, Proposition 4.33), also these sets are symplectic invariant. The same holds also for these other sets, but the proof in this case is definitely less trivial and requires a more subtle study of these action-minimizing orbits. It can be deduced, for instance, as a special case of [8, Theorem 1.10] (which also applies to the non-autonomous case). Let us denote by  $\mathcal{A}_c^*(H)$  and  $\mathcal{N}_c^*(H)$  the Aubry and Mañé sets associated to a

Tonelli Hamiltonian  $H$  (in the sense of the Legendre transform of the corresponding ones for the associated Lagrangian). Then:

**Theorem 5.37 (Bernard, [8]).** *Let  $L : TM \rightarrow \mathbb{R}$  be a Tonelli Lagrangian and  $H : T^*M \rightarrow \mathbb{R}$  the associated Hamiltonian. If  $\Phi : T^*M \rightarrow T^*M$  is an exact symplectomorphism, then*

$$\mathcal{A}_c^*(H \circ \Phi) = \Phi^{-1}(\mathcal{A}_c^*(H)) \quad \text{and} \quad \mathcal{N}_c^*(H \circ \Phi) = \Phi^{-1}(\mathcal{N}_c^*(H)).$$

This result can be easily extended to non-exact symplectomorphisms, using Lemma 4.30 (as we have already done in the case of Proposition 4.31).

**Theorem 5.38.** *Let  $L : TM \rightarrow \mathbb{R}$  be a Tonelli Lagrangian and  $H : T^*M \rightarrow \mathbb{R}$  the associated Hamiltonian. If  $\Psi : T^*M \rightarrow T^*M$  is a symplectomorphism of class  $[\Psi]$ , then*

$$\mathcal{A}_c^*(H \circ \Psi) = \Psi^{-1}(\mathcal{A}_{c+[\Psi]}^*(H)) \quad \text{and} \quad \mathcal{N}_c^*(H \circ \Psi) = \Psi^{-1}(\mathcal{N}_{c+[\Psi]}^*(H)).$$

**Remark 5.39.** More geometric proofs of this result can be obtained using weak KAM theory, for instance [54, Lemma 1] (it does not apply to all symplectomorphisms, but only to those that are Hamiltonianly isotopic to the identity) and [12]. See also [50].

## 5.B - AN EXAMPLE: THE SIMPLE PENDULUM II

In Addendum 4.B we discussed the Mather sets, the  $\alpha$ -function and the  $\beta$ -function, in the case of the special case of a simple pendulum, described by the Lagrangian:

$$\begin{aligned} L : T\mathbb{T} &\rightarrow \mathbb{R} \\ (x, v) &\mapsto \frac{1}{2}|v|^2 + (1 - \cos(2\pi x)). \end{aligned}$$

In particular we proved that:

- For all  $-\frac{4}{\pi} \leq c \leq \frac{4}{\pi}$ ,  $\widetilde{\mathcal{M}}_c = \{(0, 0)\}$ ;
- if  $c > \frac{4}{\pi}$ ,

$$\widetilde{\mathcal{M}}_{\pm c} = \{(x, v) : v = \pm \sqrt{2[(1 + \alpha(c)) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\}.$$

We want to see which are the Mañé and Aubry sets in this case. Let us start recalling the following fact, that we have proved, in a slightly different form, in Section 3.

**Proposition 5.40.** *Let  $\Lambda$  be a  $c$ -invariant (Lipschitz) Lagrangian graph in  $T^*M$ . Then, the projection on  $TM$  of each orbit on  $\Lambda$  is  $c$  semi-static.*

See for instance [24]. The proof is essentially the same as Proposition 3.8 (see also Remarks 3.9 (ii) and 5.14 (iii)). The proof extends to the case of Lipschitz  $c$ -Lagrangian graphs (i.e., Lipschitz sections that are locally the graph of closed 1-forms of cohomology class  $c$ ). Observe in fact that also in the Lipschitz case, the Hamiltonian keeps constant on invariant Lagrangian graphs (see for instance [54]).

Hence, it follows from Proposition 5.40 that for  $c > \frac{4}{\pi}$ ,

$$\widetilde{\mathcal{N}}_{\pm c} \supseteq \{(x, v) : v = \pm \sqrt{2[(1 + \alpha(\pm c)) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\} = \widetilde{\mathcal{M}}_{\pm c}.$$

In fact, it is easy to check that  $\{(x, v) : v = \pm \sqrt{2[(1 + \alpha(c)) - \cos(2\pi x)]}, \forall x \in \mathbb{T}\}$  is the graph of a closed 1-form of cohomology class  $\pm c$  (the cohomology is just the signed area enclosed between this graph and the  $x$ -axis). See also the discussion in

Addendum 4.B.

Moreover,  $\tilde{\mathcal{N}}_{\pm c}$  must be equal to  $\widetilde{\mathcal{M}}_{\pm c}$ , since it is connected (Corollary 5.34) and it must be contained in the energy level corresponding to the value  $\alpha(\pm c)$ . Therefore, recalling the inclusions in Theorem 5.21, we can conclude that

$$\tilde{\mathcal{N}}_{\pm c} = \tilde{\mathcal{A}}_{\pm c} = \widetilde{\mathcal{M}}_{\pm c} \quad \text{for all } c > \frac{4}{\pi}.$$

Let us see what happens for  $|c| \leq \frac{4}{\pi}$ . Observe that they all correspond to the same energy level, namely the one of the separatrices ( $\alpha(c) = 0$  in this case). In this energy level there are exactly three orbits:

- The fixed point  $(0, 0)$ , *i.e.*,  $\gamma_0(t) \equiv 0$ .
- The upper separatrix  $\gamma_+$ ; for instance let us choose the parametrization given by

$$\gamma_+(t) = \pi \left( \Phi_t^L(1/2, 2) \right).$$

- The lower separatrix  $\gamma_-$ ; for instance let us choose the parametrization given by

$$\gamma_-(t) = \pi \left( \Phi_t^L(1/2, -2) \right).$$

Observe that because of the symmetry of  $L$  and the chosen parametrizations, we have that  $\gamma_+(t) = \gamma_-(-t)$  for all  $t \in \mathbb{R}$ .

First of all, let us show that for  $|c| < \frac{4}{\pi}$ , neither  $\gamma_+$  or  $\gamma_-$  can be  $c$  semi-static. In fact, let us consider a  $|c|$ -closed 1-form  $\eta_c$  whose graph is contained in the region between the separatrices. This is possible since  $c$  is less than  $4/\pi$  and the cohomology represents the signed area of the region between the curve and the  $x$ -axis. Moreover, since  $|c|$  is strictly less than  $4/\pi$ , there will be a positive measure subset of  $\mathbb{T}$  on which  $H(x, \eta_c(x)) < 0$ .

If  $\gamma_+$  were  $c$ -semistatic (similarly for  $\gamma_-$ ), then using that  $\gamma_+$  is asymptotic in the past and in the future to 0, that  $\alpha(c) = 0$  and that the Mañé potential  $\phi_{\eta_c, 0}$  is Lipschitz continuous (Proposition 5.9 (3)), we obtain:

$$\begin{aligned} \phi_{\eta_c, 0}(0, 0) &= \lim_{T \rightarrow +\infty} \phi_{\eta_c, 0}(\gamma_+(-T), \gamma_+(T)) = \\ &= \lim_{T \rightarrow +\infty} \int_{-T}^T L_{\eta_c}(\gamma_+(t), \dot{\gamma}_+(t)) dt = \\ &\geq \lim_{T \rightarrow +\infty} \int_{-T}^T (\eta_c(\gamma_+(t)) \cdot \dot{\gamma}_+(t) - H(\gamma_+(t), \eta_c(\gamma_+(t)))) dt = \\ &= - \lim_{T \rightarrow +\infty} \int_{-T}^T H(\gamma_+(t), \eta_c(\gamma_+(t))) dt > 0 \end{aligned}$$

where the third inequality comes from Fenchel-Legendre inequality and the last one from the fact that there exists a positive measure set of  $\mathbb{T}$  in which  $H(\gamma_+(t), \eta_c(\gamma_+(t)))$  is strictly negative. But this leads to a contradiction, since the action of the constant path  $\gamma_0$  is zero:  $A_{L_{\eta_c}}(\gamma_0) = 0$ .

We have just proved that (use also Theorem 5.21):

$$\tilde{\mathcal{N}}_c = \tilde{\mathcal{A}}_c = \widetilde{\mathcal{M}}_c = \{(0, 0)\} \quad \text{for all } |c| < \frac{4}{\pi}.$$

Finally, let us consider the case  $c_{\pm} := \pm \frac{4}{\pi}$ . As above, it follows from Proposition 5.40 (taking the graphs of the separatrices as invariant Lipschitz Lagrangian graphs) that:

$$\tilde{\mathcal{N}}_{c_{\pm}} = \{(x, v) : v = \pm \sqrt{2[1 - \cos(2\pi x)]}, \forall x \in \mathbb{T}\} \supset \widetilde{\mathcal{M}}_{c_{\pm}}.$$

We want to show that also in this case  $\tilde{\mathcal{A}}_{c_{\pm}} = \tilde{\mathcal{N}}_{c_{\pm}}$  (but observe that this time they contain the Mather set properly). Let us denote by  $\eta_{\pm}$  the closed 1-forms given



by the graphs of (respectively) the upper and lower separatrix. We have already pointed out that  $[\eta_{\pm}] = c_{\pm}$ . The key observation is the following lemma.

**Lemma 5.41.** *For every  $x, y \in \mathbb{T}$  we have  $\phi_{\eta_+,0}(x, y) = -\phi_{\eta_-,0}(y, x)$ .*

Observe that it follows immediately from this Lemma that the upper (resp. lower) separatrix is not only  $c_+$  semi-static (resp.  $c_-$  semi-static), but it is indeed  $c_+$  static (resp.  $c_-$  static). Therefore,  $\tilde{\mathcal{N}}_{c_{\pm}} = \tilde{\mathcal{A}}_{c_{\pm}}$ .

**Proof.** The above equality is always true if  $x = y$  (see Proposition 5.9 (4)). Suppose without any loss of generality that  $x < y$  and let  $S < T$  such that  $\gamma_+(S) = x = \gamma_-(-S)$  and  $\gamma_+(T) = y = \gamma_-(-T)$  (this is true only for the parametrization that we chose above). Now, using that these curves are semi-static (for their respective cohomologies), the symmetry of  $L$ , the relation between  $\gamma_+$  and  $\gamma_-$  and the fact that  $\eta_+ = -\eta_-$ , we obtain:

$$\begin{aligned} \phi_{\eta_+,0}(x, y) &= \int_S^T (L(\gamma_+(t), \dot{\gamma}_+(t)) - \eta_+(\gamma_+(t)) \cdot \dot{\gamma}_+(t)) dt = \\ &= \int_S^T (L(\gamma_-(-t), -\dot{\gamma}_-(-t)) + \eta_-(\gamma_-(t)) \cdot (-\dot{\gamma}_-(-t))) dt = \\ &= - \int_{-T}^{-S} (L(\gamma_-(s), \dot{\gamma}_-(s)) - \eta_-(\gamma_-(s)) \cdot \dot{\gamma}_-(s)) dt = \\ &= -\phi_{\eta_-,0}(y, x). \end{aligned}$$

□

Summarising:

- For  $c > \frac{4}{\pi}$ :

$$\tilde{\mathcal{N}}_{\pm c} = \tilde{\mathcal{A}}_{\pm c} = \tilde{\mathcal{M}}_{\pm c} = \pm \sqrt{2[(1 + \alpha(c)) - \cos(2\pi x)]}, \forall x \in \mathbb{T}.$$

- For  $|c| < \frac{4}{\pi}$ :

$$\tilde{\mathcal{N}}_c = \tilde{\mathcal{A}}_c = \tilde{\mathcal{M}}_c = \{(0, 0)\}.$$

- For  $c_{\pm} = \pm \frac{4}{\pi}$ :

$$\tilde{\mathcal{N}}_{c_{\pm}} = \tilde{\mathcal{A}}_{c_{\pm}} = \{(x, v) : v = \pm \sqrt{2[1 - \cos(2\pi x)]}, \forall x \in \mathbb{T}\} \supset \tilde{\mathcal{M}}_{c_{\pm}} = \{(0, 0)\}.$$

**Remark 5.42.** (i) For  $c_{\pm} = \pm \frac{4}{\pi}$  we have examples in which the Mather set is strictly contained in the Aubry and Mañé sets.

(ii) In all the above examples, the Aubry set coincides with the Mañé set. However it is possible to find examples in which this is not anymore true. For instance, consider the double covering  $2\mathbb{T}$  of  $\mathbb{T}$  and let  $\rho : 2\mathbb{T} \rightarrow \mathbb{T}$  be the covering map. Let us lift the Lagrangian of the simple pendulum to a Lagrangian  $\tilde{L}$  on  $\mathbb{T}(2\mathbb{T})$ , given by  $\tilde{L}(\tilde{x}, v) = L(\rho(\tilde{x}), v)$ . Observe that one can equivalently consider the system on  $\mathbb{T}\mathbb{T}$  given by  $L(x, v) = \frac{1}{2}|v|^2 - (\cos(4\pi x) - 1)$ . One can verify that (we specify the dependence on the Lagrangian and consider the projected sets):

- $\mathcal{M}_0(\tilde{L}) = \rho^{-1}(\mathcal{M}_0(L)) = \rho^{-1}(\{0\}) = \{0, 1/2\}$ .
- $\mathcal{A}_0(\tilde{L}) = \rho^{-1}(\mathcal{A}_0(L)) = \rho^{-1}(\{0\}) = \{0, 1/2\}$ .

This result is true in general when we consider a finite covering (see [17, Lemma 2.3]). However, the same is not true anymore for the Mañé set. In fact, one can check that the four separatrices connecting the two minimizing fixed points are all 0-semistatic and therefore:

$$\mathcal{N}_0(\tilde{L}) = 2\mathbb{T} \supset \{0, \frac{1}{2}\} = \rho^{-1}(\mathcal{N}_0(L)).$$

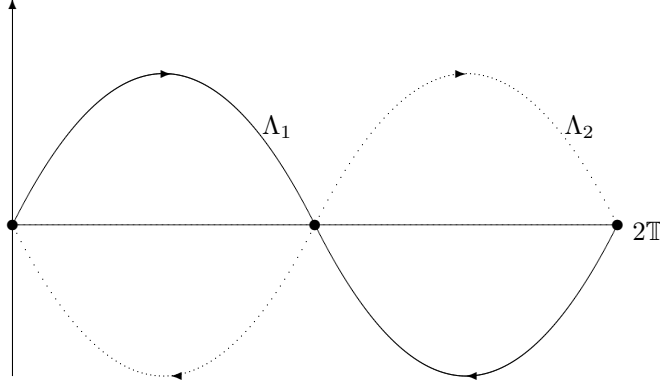


FIGURE 3. The two invariant Lagrangian graphs  $\Lambda_1$  (continuous line) and  $\Lambda_2$  (dashed line) for the “lifted” pendulum on  $2\mathbb{T}$ .

In fact, it is easy to verify that the lifted system has two exact invariant Lagrangian graphs (zero-cohomology = zero area) made by the lifts of the separatrices (combined so to enclose zero area). Therefore, the lifts of the separatrices are 0 semi-static (since they are contained on 0-Lagrangian graphs), but they are not 0-static (see figure 3).

### 5.C - MATHER’S APPROACH: PEIERLS’ BARRIER

In this addendum we want to describe Mather’s original approach to the study of action-minimizing curves. As far as the definition of the Mañé set is concerned, it is pretty much the same as the one we have described before. The main difference is in the definition of the Aubry set. In fact, Mather’s original definition considered what he called *c-regular-minimizers*. We shall see that they indeed coincide with *c-static* curves.

The main ingredient in Mather’s approach is the notion of *Peierls Barrier*, introduced in [44].<sup>2</sup>

For  $t > 0$  and  $x, y \in M$ , let us consider :

$$(24) \quad h_{\eta,t}(x, y) = \min \int_0^t L_{\eta}(\gamma(s), \dot{\gamma}(s)) ds ,$$

where the minimum is taken over all piecewise  $C^1$  paths  $\gamma : [0, t] \rightarrow M$ , such that  $\gamma(0) = x$  and  $\gamma(t) = y$ . This minimum is achieved because of Tonelli theorem (Theorem 5.1). We define the *Peierls barrier* as:

$$(25) \quad h_{\eta}(x, y) = \liminf_{t \rightarrow +\infty} (h_{\eta,t}(x, y) + \alpha(c)t) .$$

**Remark 5.43.** (i) Observe that  $h_{\eta}$  does not depend only on the cohomology class  $c$ , but also on the choice of the representative  $\eta$ ; namely, if  $\eta' = \eta + df$ , then  $h_{\eta'}(x, y) = h_{\eta}(x, y) + f(y) - f(x)$ . Anyhow, this dependence will not be harmful for what we are going to do in the following (it will not change the set of action-minimizing curves).

(ii) The main difference between Peierls barrier and Mañé potential is that in this

<sup>2</sup>The function that we are defining here is actually a slight generalization of  $h_c^{\infty}$  defined in [44]. Pay attention that throughout this article, the sign of the  $\alpha$  function is wrong: wherever there is  $\alpha(c)$ , it should be substituted by  $-\alpha(c)$ .

case we consider curves defined over longer and longer time intervals. In particular, it is easy to check that:

$$h_\eta(x, y) \geq \phi_{\eta, \alpha(c)}(x, y) \quad \forall x, y \in M.$$

(iii) This function  $h_\eta$  is a generalization of Peierls Barrier introduced by Aubry [3] and Mather [39, 41, 43] in their study of twist maps. In some sense we are comparing, in the limit, the action of Tonelli minimizers of time length  $T$  with the corresponding average  $c$ -minimal action  $-\alpha(c)T$ . Remember, in fact, that  $-\alpha(c)$  is the “average action” of a  $c$ -minimal measure.

(iv) Albert Fathi [22] showed that - in the autonomous case - this  $\liminf$  can be replaced with a  $\lim$ . This is not generally true in the non-autonomous time-periodic case (see for instance [25] for some counterexamples); Tonelli Lagrangians for which this convergence result holds are called *regular*. Patrick Bernard [5] showed that under suitable assumptions on the Mather set it is possible to prove that the Lagrangian is regular. For instance, if the Mather set  $\tilde{\mathcal{M}}_c$  is union of 1-periodic orbits, then  $L_\eta$  is regular. This problem turned out to be strictly related to the convergence of the so-called *Lax-Oleinik semigroup* (see [22] for its definition).

Analogously to Proposition 5.9, one can prove the following.

**Proposition 5.44.** *The values of the map  $h_\eta$  are finite. Moreover, the following properties hold:*

i) *for each  $x, y, z \in M$  and  $t > 0$*

$$h_\eta(x, y) \leq h_\eta(x, z) + h_{\eta, t}(z, y) + \alpha(c)t;$$

ii) *for each  $x, y, z \in M$ ,  $h_\eta(x, y) \leq h_\eta(x, z) + h_\eta(z, y)$ .*

iii) *for each  $x, y, z \in M$ ,  $h_\eta(x, y) \leq h_\eta(x, z) + \phi_{\eta, \alpha(c)}(z, y)$ .*

iv) *for each  $x, y, z \in M$  and  $t > 0$*

$$h_\eta(x, y) \leq h_\eta(x, z) + h_{\eta, t}(z, y) + \alpha(c)t;$$

v) *for each  $x \in M$ ,  $h_\eta(x, x) \geq 0$ ;*

vi) *for each  $x, y \in M$ ,  $h_\eta(x, y) + h_\eta(y, x) \geq 0$ .*

It is interesting to consider the following symmetrization:

$$(26) \quad \begin{aligned} \delta_c : M \times M &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto h_\eta(x, y) + h_\eta(y, x). \end{aligned}$$

Observe that this function does now depend only on the cohomology class  $c$  and moreover it is non-negative, symmetric and satisfies the triangle inequality.

An interesting property of  $\delta_c$  is the following (see [44, Section 8]). If  $d$  denotes the distance induced on  $M$  by the Riemannian metric  $g$ , then there exists  $C > 0$  such that for each  $x, y \in M$  we have

$$\delta_c(x, y) \leq Cd(x, y)^2.$$

**Remark 5.45.** The same estimate continues to be true for the non-autonomous time-periodic case. In this case we have that

$$\delta_c((x, \tau_0), (y, \tau_1)) \leq C[d(x, y) + \|\tau_1 - \tau_0\|]^2$$

for each  $(x, \tau_0), (y, \tau_1) \in \mathcal{A}_c$ , where

$$\|\tau_1 - \tau_0\| = \inf \{|t_1 - t_0| : t_i \in \mathbb{R}, t_i \equiv \tau_i \pmod{1}, i = 0, 1\}.$$

Let us see now some relation between this Peierls barrier (or equivalently  $\delta_c$ ) and  $c$ -action minimizing curves. Let  $\gamma : \mathbb{R} \longrightarrow M$  be a  $c$ -minimizer and consider

$x_\alpha, x'_\alpha$  in the  $\alpha$ -limit set<sup>3</sup> of  $\gamma$  and  $x_\omega, x'_\omega$  in the  $\omega$ -limit set<sup>4</sup> of  $\gamma$ . John Mather in [44, Section 6] proved that  $\delta_c(x_\alpha, x'_\alpha) = \delta_c(x_\omega, x'_\omega) = 0$ . In general, it is not true that  $\delta_c(x_\alpha, x_\omega) = 0$ ; what one can prove is that this value does not depend on the particular  $x_\alpha$  and  $x_\omega$ , i.e.,  $\delta_c(x_\alpha, x_\omega) = \delta_c(x'_\alpha, x'_\omega)$ : it is a property of the limit sets rather than of their elements. Nevertheless, there will exist particular  $c$ -minimizers for which this value is equal to 0 and these will be the  $c$ -minimizers that we want to single out.

**Definition 5.46 (c-regular minimizers).** *A  $c$ -minimizer  $\gamma : \mathbb{R} \rightarrow M$  is called a  $c$ -regular minimizer, if  $\delta_c(x_\alpha, x_\omega) = 0$  for each  $x_\alpha$  in the  $\alpha$ -limit set of  $\gamma$  and  $x_\omega$  in the  $\omega$ -limit set of  $\gamma$ .*

Mather defined the *Aubry set* as the union of the support of all these  $c$ -regular minimizers.

**Definition 5.47 (Aubry set).** *The Aubry set (with cohomology class  $c$ ) is:*

$$\tilde{\mathcal{A}}_c = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \gamma \text{ is a } c\text{-regular minimizer and } t \in \mathbb{R}\}.$$

it turns out that this set coincides exactly with the one that we have defined in Section 5. In fact, one can prove that:

**Proposition 5.48.**  *$\gamma$  is a  $c$ -regular minimizer of  $L$  if and only if  $\gamma$  is a  $c$ -static curve of  $L$ .*

A proof of this can be found in [22, Proposition 9.2.5]. The essential ingredient is that  $\phi_{\eta, \alpha(c)}(x, y) = h_\eta(x, y)$  if  $x, y \in \mathcal{A}_c$  (see also Remark 6.28).

Moreover, one can also provide another alternative definition of the (projected) Aubry set:

**Proposition 5.49** (See [22, Proposition 5.3.8]). *The following properties are equivalent.*

- i)  $x \in \mathcal{A}_c$ ;
- ii)  $h_\eta(x, x) = 0$ ;
- iii) *there exists a sequence of absolutely continuous curves  $\gamma_n : [0, t_n] \rightarrow M$  such that:*
  - for each  $n$ , we have  $\gamma_n(0) = \gamma_n(t_n) = x$ ;
  - the sequence  $t_n \rightarrow +\infty$ , as  $n \rightarrow +\infty$ ;
  - as  $n \rightarrow +\infty$ ,  $\int_0^{t_n} L_\eta(\gamma_n(s), \dot{\gamma}_n(s)) ds + \alpha(c)t_n \rightarrow 0$ .

**Remark 5.50.** (i) Therefore, the Aubry set consists of points that are contained in loops with period as long as we want and action as close as we want to the minimal average one.

(ii) Moreover, it follows from ii) in Proposition 5.49 that  $\delta_c$  is a pseudometric on the projected Aubry set

$$\mathcal{A}_c = \{x \in M : \delta_c(x, x) = 0\}.$$

(iii) One can easily construct a metric space out of  $(\mathcal{A}_c, \delta_c)$ . We call *quotient Aubry set* the metric space  $(\bar{\mathcal{A}}_c, \bar{\delta}_c)$  obtained by identifying two points in  $\mathcal{A}_c$ , if their  $\delta_c$ -pseudodistance is zero. This set plays quite an interesting role in the study of the dynamics; see for example [46, 47, 53, 23] for more details.

<sup>3</sup> Recall that a point  $z$  is in the  $\alpha$ -limit set of  $\gamma$ , if there exists a sequence  $t_n \rightarrow -\infty$  such that  $\gamma(t_n) \rightarrow z$ .

<sup>4</sup> Recall that a point  $z$  is in the  $\omega$ -limit set of  $\gamma$ , if there exists a sequence  $t_n \rightarrow +\infty$  such that  $\gamma(t_n) \rightarrow z$ .

## 6. WEAK KAM THEORY

Another interesting approach to the study of these invariant sets is provided by the so-called *weak KAM theory*, which represents the functional analytical counterpart of the variational methods discussed in the previous sections. In section 3 in fact, we pointed out the relation between KAM tori (or more generally, invariant Lagrangian graphs) and classical solutions and subsolutions of Hamilton-Jacobi equation (see Remark 3.2 (iv)). This approach that we are going to describe, will be based on studying “weak” (non-classical) solutions of Hamilton-Jacobi equation and some special class of subsolutions (*critical subsolutions*). From a more geometrical point of view, this can be interpreted as the study of particular Lagrangian graphs (not necessarily invariant) and their non-removable intersection (see also [50]). This point of view makes this approach particularly interesting, since it relates the dynamics of the system to the geometry of the space and might potentially open the way to a “symplectic” definition of Aubry-Mather theory (see also Sections 4A, 5A and [8, 54, 12]).

In this section we want to provide a brief presentation of this theory, omitting most of the proofs, for which we refer the reader to the excellent - and self-contained - presentation [22].

The main object of investigation is represented by Hamilton-Jacobi (H-J) equation:

$$H_\eta(x, d_x u) = H(x, \eta(x) + d_x u) = k,$$

where  $\eta$  is a closed 1-form on  $M$  with a certain cohomology class  $c$ . Observe that considering H-J equations for different 1-forms corresponding to different cohomology classes, is equivalent to Mather's idea of changing Lagrangian (see section 4).

From now on, we shall consider  $L$  to be a Tonelli Lagrangian on a compact manifold  $M$  and  $H$  its associated Hamiltonian. Let us fix  $\eta$  to be a closed 1-form on  $M$  with cohomology class  $c$ , and, as before, denote by  $L_\eta$  and  $H_\eta$  the modified Lagrangian and Hamiltonian. In classical mechanics, one is interested in studying solutions of this equation, *i.e.*,  $C^1$  functions  $u : M \rightarrow \mathbb{R}$  such that  $H_\eta(x, d_x u) = k$ . It is immediate to check that for any given cohomology class there exists at most one value of  $k$  for which these  $C^1$  solutions may exist. In fact, it is enough to observe that if  $u$  and  $v$  are two  $C^1$  functions on a compact manifold, there will exist a point  $x_0$  at which their differentials coincide (take any critical point of  $u - v$ ). We shall see (Theorem 6.16) that this value of  $k$  for which solutions may exist, coincides with  $\alpha(c)$  or Mañé critical value  $c(L_\eta)$  (defined in sections 4 and 5).

**Remark 6.1.** The existence of such solutions has significant implications to the dynamics of the system and it is, consequently, quite rare. In particular, they correspond to Lagrangian graphs, which are invariant under the Hamiltonian flow  $\Phi_t^H$  (Hamilton-Jacobi theorem). For instance, in the case of  $M = \mathbb{T}^d$  and nearly-integrable systems these solutions correspond to KAM tori (this might give an idea of their rareness).

One of the main results of weak KAM theory is that, in the case of Tonelli Hamiltonians, a weaker kind of solutions do always exist. In the following we are going to define these generalized solutions and their relation with the dynamics of the system. It is important to point out that one of the main ingredient in the proof of all these results is provided by Fenchel inequality (cf. 3 in Section 2).

Let us start by generalizing the concept of subsolution. In the  $C^1$ -case it is easy to check - using Fenchel inequality - that the following property holds (the proof is essentially the same as for (8) in Section 3).

**Proposition 6.2.** *Let  $u : M \rightarrow \mathbb{R}$  be  $C^1$ ;  $u$  satisfies  $H(x, \eta(x) + d_x u) \leq k$  for all  $x \in M$  if and only if for all  $a < b$  and  $\gamma : [a, b] \rightarrow M$*

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + k(b - a).$$

This last inequality provides the ground for to a definition of subsolution in the  $C^0$ -case.

**Definition 6.3 (Dominated functions).** *Let  $u : M \rightarrow \mathbb{R}$  be a continuous function;  $u$  is dominated by  $L_\eta + k$ , and we shall write  $u \prec L_\eta + k$ , if for all  $a < b$  and  $\gamma : [a, b] \rightarrow M$*

$$(27) \quad u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + k(b - a).$$

One can check that if  $u \prec L_\eta + k$  then  $u$  is Lipschitz and its Lipschitz constant can be bounded by a constant  $C(k)$  independent of  $u$ ; in fact, it is sufficient to apply the definition of dominated function with the speed-one geodesic connecting any two points  $x$  and  $y$  and consider the maximum of  $L$  over the unit tangent ball (see [22, Proposition 4.2.1 (iii)]). In particular, all dominated functions for values of  $k$  in a compact set are equiLipschitz. On the other hand, it is easy to check that each Lipschitz function is dominated by  $L_\eta + k$ , for a suitable  $k$  depending on its Lipschitz constant: this shows that dominated functions exist.

Dominated functions generalize subsolutions of H-J to the continuous case. In fact:

**Proposition 6.4** (see [22, Theorem 4.25]). *If  $u \prec L_\eta + k$  and  $d_x u$  exists, then  $H(x, \eta(x) + d_x u) \leq k$ . Moreover, if  $u : M \rightarrow \mathbb{R}$  is Lipschitz and  $H(x, \eta(x) + d_x u) \leq k$  a.e., then  $u \prec L_\eta + k$ .*

**Remark 6.5.** Using the fact that any Lipschitz function is differentiable almost everywhere (Rademacher theorem), one could equivalently define subsolutions in the following way: *a locally Lipschitz function  $u : M \rightarrow \mathbb{R}$  is a subsolution of  $H_\eta(x, d_x u) = k$ , with  $k \in \mathbb{R}$ , if  $H_\eta(x, d_x u) \leq k$  for almost every  $x \in M$ .*

**Remark 6.6.** One interesting question is: for which values of  $k$  do there exist functions dominated by  $L_\eta + k$  (or equivalently subsolutions of  $H(x, \eta(x) + d_x u) = k$ )? It is possible to show that there exists a value  $k_c \in \mathbb{R}$  such that  $H(x, \eta(x) + d_x u) = k$  does not admit any subsolution for  $k < k_c$ , while it has subsolutions for  $k \geq k_c$ , see [32, 22]. In particular, if  $k > k_c$  there exist  $C^\infty$  subsolutions. It turns out that the constant  $k_c$  coincides with  $\alpha(c)$  and the Mañé's critical value (where  $c = [\eta]$ ). See [15, 22].

Functions corresponding to this “critical domination” play an important role, since they encode significant information about the dynamics of the system.

**Definition 6.7 (Critical subsolutions).** *A function  $u \prec L_\eta + \alpha(c)$  is said to be critically dominated. Equivalently, we shall also call it an  $\eta$ -critical subsolution, since  $H(x, \eta(x) + d_x u) \leq \alpha(c)$  for almost every  $x \in M$ .*

**Remark 6.8.** The above observation provides a further definition of  $\alpha(c)$ :

$$\alpha(c) = \inf_{u \in C^\infty(M)} \max_{x \in M} H(x, \eta(x) + d_x u).$$

This theorem has been proven in [15, Theorem A] (see also in [22]). This infimum is not a minimum, but it becomes a minimum over the set of Lipschitz functions on  $M$  (also over the smaller set of  $C^{1,1}$  functions, see the addendum at the end of this section and [26, 6]). This characterization has the following geometric interpretation. If we consider the space  $T^*M$  equipped with the canonical symplectic form, the

graph of the differential of a  $C^1$   $\eta$ -critical subsolution (plus the 1-form  $\eta$ ) is nothing else than a  $c$ -Lagrangian graph (*i.e.*, a Lagrangian graph with cohomology class  $c$ ). Therefore Mañé  $c$ -critical energy level  $\mathcal{E}_c^* = \{(x, p) \in T^*M : H(x, p) = \alpha(c)\}$  corresponds to a  $(2d - 1)$ -dimensional hypersurface, such that the region it bounds is convex in each fiber and does not contain in its interior any  $c$ -Lagrangian graph, while any of its neighborhoods does.

Analogously to what we have already seen for subsolutions, it would be interesting to investigate if there existed an equivalent characterization of classical solutions of Hamilton-Jacobi, that does not involve the regularity of the solution. This would allow us to define “*weak*” solutions, that hopefully are not so rare as the classical ones.

Let us now recall some properties of classical solutions, which will allow us to provide a “weaker” definition of solution (the proof of this proposition is essentially the same as for (8) in Section 3).

**Proposition 6.9** (see [22, Theorem 4.1.10]). *Let  $u : M \rightarrow \mathbb{R}$  a  $C^1$  function and  $k \in \mathbb{R}$ . The following conditions are equivalent:*

- (1)  *$u$  is solution of  $H(x, \eta(x) + d_x u) = k$ ;*
- (2)  *$u \prec L_\eta + k$  and for each  $x \in M$  there exists  $\gamma_x : (-\infty, +\infty) \rightarrow M$  such that  $\gamma_x(0) = x$  and for any  $[a, b]$ :*

$$u(\gamma_x(b)) - u(\gamma_x(a)) = \int_a^b L_\eta(\gamma_x(t), \dot{\gamma}_x(t)) dt + k(b - a).$$

- (3)  *$u \prec L_\eta + k$  and for each  $x \in M$  there exists  $\gamma_x : (-\infty, 0] \rightarrow M$  such that  $\gamma_x(0) = x$  and for any  $a < b \leq 0$ :*

$$u(\gamma_x(b)) - u(\gamma_x(a)) = \int_a^b L_\eta(\gamma_x(t), \dot{\gamma}_x(t)) dt + k(b - a).$$

- (4)  *$u \prec L_\eta + k$  and for each  $x \in M$  there exists  $\gamma_x : [0, +\infty) \rightarrow M$  such that  $\gamma_x(0) = x$  and for any  $0 \leq a < b$ :*

$$u(\gamma_x(b)) - u(\gamma_x(a)) = \int_a^b L_\eta(\gamma_x(t), \dot{\gamma}_x(t)) dt + k(b - a).$$

Inspired by this fact, let us consider the curves for which equality in (27) holds.

**Definition 6.10 (Calibrated curves).** *Let  $u \prec L_\eta + k$ . A curve  $\gamma : I \rightarrow M$  is  $(u, L_\eta, k)$ -calibrated if for any  $[a, b] \subseteq I$*

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt + k(b - a).$$

These curves are very special curves and it turns out that they are orbits of the Euler-Lagrange flow. In fact:

**Proposition 6.11.** *If  $u \prec L_\eta + k$  and  $\gamma : [a, b] \rightarrow M$  is  $(u, L_\eta, k)$ -calibrated, then  $\gamma$  is a  $c$ -Tonelli minimizer, *i.e.*,*

$$\int_a^b L_\eta(\gamma(t), \dot{\gamma}(t)) dt \leq \int_a^b L_\eta(\sigma(t), \dot{\sigma}(t)) dt$$

*for any  $\sigma : [a, b] \rightarrow M$  such that  $\sigma(a) = \gamma(a)$  and  $\sigma(b) = \gamma(b)$ . Most of all, this implies that  $\gamma$  is a solution of the Euler-Lagrange flow and therefore it is  $C^r$  (if  $L$  is  $C^r$ ).*

The proof of this result is the same as the one of Proposition 3.8 (see also [22, Proposition 4.3.2 and Corollary 4.3.3]).

Moreover, the following differentiability result holds.

**Proposition 6.12** (see [22, Theorem 4.3.8]). *Let  $u \prec L_\eta + k$  and  $\gamma : [a, b] \rightarrow M$  be  $(u, L_\eta, k)$ -calibrated.*

- i) *If  $d_{\gamma(t)}u$  exists for some  $t \in [a, b]$ , then  $H(\gamma(t), \eta(\gamma(t))) + d_{\gamma(t)}u = k$  and  $d_{\gamma(t)}u = \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t))$ .*
- ii) *If  $t \in (a, b)$ , then  $d_{\gamma(t)}u$  exists.*

**Remark 6.13.** Calibrated curves are “Lagrangian gradient lines” of  $\text{grad}_L u$  (where  $\text{grad}_L u$  is a multivalued vector field given by the equation  $d_x u = \frac{\partial L}{\partial v}(x, \text{grad}_L u)$ ). Therefore, there is only one possibility for calibrated curves, at each point of differentiability of  $u$ .

This suggests the following definitions.

**Definition 6.14 (weak KAM solutions).** *Let  $u \prec L_\eta + k$ .*

- *$u$  is a weak KAM solution of negative type (or backward Weak KAM solution) if for each  $x \in M$  there exists  $\gamma_x : (-\infty, 0] \rightarrow M$  such that  $\gamma_x(0) = x$  and  $\gamma_x$  is  $(u, L_\eta, k)$ -calibrated;*
- *$u$  is a weak KAM solution of positive type (or forward Weak KAM solution) if for each  $x \in M$  there exists  $\gamma_x : [0, +\infty) \rightarrow M$  such that  $\gamma_x(0) = x$  and  $\gamma_x$  is  $(u, L_\eta, k)$ -calibrated.*

**Remark 6.15.** Observe that any weak KAM solution of negative type  $u_-$  (resp. of positive type  $u_+$ ) for a given Lagrangian  $L$ , can be seen as a weak KAM solution of positive type (resp. of negative type) for the *symmetrical* Lagrangian  $\tilde{L}(x, v) := L(x, -v)$ .

Let us denote with  $\mathcal{S}_\eta^-$  the set of Weak KAM solutions of negative type and  $\mathcal{S}_\eta^+$  the ones of positive types. Albert Fathi [19, 22] proved that these sets are always non-empty.

**Theorem 6.16 (Weak KAM theorem).** *There is only one value of  $k$  for which weak KAM solutions of positive or negative type of  $H(x, \eta(x) + d_x u) = k$  exist. This value coincides with  $\alpha(c)$ , where  $\alpha : H^1(M; \mathbb{R}) \rightarrow \mathbb{R}$  is Mather’s  $\alpha$ -function. In particular, for any  $u \prec L_\eta + \alpha(c)$  there exist a weak KAM solution of negative type  $u_-$  and a weak KAM solution of positive type  $u_+$ , such that  $u_- = u = u_+$  on the projected Aubry set  $\mathcal{A}_c$ .*

Therefore, for any given weak KAM solution of negative type  $u_-$  (resp. of positive type  $u_+$ ), there exists a weak KAM solution of positive type of positive type  $u_+$  (resp. of negative type  $u_-$ ) such that  $u_- = u_+$  on the projected Aubry set  $\mathcal{A}_c$ . In particular:

**Proposition 6.17** (see [22, Theorem 4.12.6]). *The projected Mather set  $\mathcal{M}_c$  is the uniqueness set for weak KAM solutions of the same type. Namely, if  $u_-, v_-$  are weak KAM solutions of negative type (resp.  $u_+, v_+$  are weak KAM solutions of positive type) and  $u_- = v_-$  on  $\mathcal{M}_c$  (resp.  $u_+ = v_+$  on  $\mathcal{M}_c$ ), then they coincide everywhere on  $M$ .*

Two solutions  $u_-$  and  $u_+$  that coincide on the (projected) Mather set are said to be *conjugate*. We shall denote by  $(u_-, u_+)$  a couple of conjugate subsolutions.



Let us try to understand the dynamical meaning of such solutions. Albert Fathi [22] - using these generalized solutions - proved a *weak* version of Hamilton-Jacobi theorem, showing the relation between these weak solutions and the dynamics of the associated Hamiltonian system. We shall state it for weak KAM solution of negative type, but - using remark 6.15 - one can deduce an analogous statement for weak KAM solutions of positive type.

**Theorem 6.18 (Weak Hamilton-Jacobi Theorem, [22, Theorem 4.13.2]).** *Let  $u_- : M \rightarrow \mathbb{R}$  be a weak KAM solution of negative type and consider*

$$\text{Graph}(\eta + du_-) := \{(x, \eta(x) + d_x u_-), \text{ where } d_x u_- \text{ exists}\}.$$

*Then:*

- i)  $\overline{\text{Graph}(\eta + du_-)}$  is compact and is contained in the energy level  $\mathcal{E}_c^* = \{(x, p) \in T^*M : H(x, p) = \alpha(c)\}$ ;
- ii)  $\Phi_{-t}^H(\overline{\text{Graph}(\eta + du_-)}) \subseteq \text{Graph}(\eta + du_-)$  for each  $t > 0$ ;
- iii)  $M = \pi(\overline{\text{Graph}(\eta + du_-)})$ , where  $\pi : T^*M \rightarrow M$  is the canonical projection.

Moreover, let us define:

$$\mathcal{I}^*(u_-) := \bigcap_{t \geq 0} \Phi_{-t}^H(\overline{\text{Graph}(\eta + du_-)}).$$

$\mathcal{I}^*(u_-)$  is non-empty, compact and invariant under  $\Phi_t^H$ . Furthermore, its “unstable set” contains  $\overline{\text{Graph}(\eta + du_-)}$ ; i.e.,

$$\overline{\text{Graph}(\eta + du_-)} \subseteq W^u(\mathcal{I}^*(u_-)) := \left\{ (y, p) : \text{dist}(\Phi_{-t}^H(y, p), \mathcal{I}^*(u_-)) \xrightarrow{t \rightarrow +\infty} 0 \right\}.$$

There is a relation between these invariant sets  $\mathcal{I}^*(u_-)$  (or  $\mathcal{I}^*(u_+)$ ) and the Aubry set  $\tilde{\mathcal{A}}_c$  (recall that  $\mathcal{I}^*(u_-), \mathcal{I}^*(u_+) \subset T^*M$ , while  $\tilde{\mathcal{A}}_c \subset TM$ ).

**Theorem 6.19.**

$$\begin{aligned} \mathcal{A}_c^* := \mathcal{L}(\tilde{\mathcal{A}}_c) &= \bigcap_{u_- \in \mathcal{S}_\eta^-} \mathcal{I}^*(u_-) = \bigcap_{u_- \in \mathcal{S}_\eta^-} \text{Graph}(\eta + du_-) = \\ &= \bigcap_{u_+ \in \mathcal{S}_\eta^+} \mathcal{I}^*(u_+) = \bigcap_{u_+ \in \mathcal{S}_\eta^+} \text{Graph}(\eta + du_+) = \\ (28) \quad &= \bigcap_{(u_-, u_+)} \{x \in M : u_-(x) = u_+(x)\}, \end{aligned}$$

where  $\mathcal{L} : TM \rightarrow T^*M$  denotes the Legendre transform of  $L$  and  $(u_-, u_+)$  are conjugates solutions.

Since it is easier to work with subsolutions rather than weak solutions, we want to discuss now how  $\eta$ -critical subsolutions, although they contain less dynamical information than weak KAM solutions, can be used to characterize Aubry and Mañé sets in a similar way.

Consider  $u \prec L_\eta + \alpha(c)$ . For  $t \geq 0$  define

$$\tilde{\mathcal{I}}_t(u) := \{(x, v) \in TM : \gamma_{(x, v)}(s) := \pi \Phi_s^L((x, v)) \text{ is } (u, L_\eta, \alpha(c)) - \text{calibr. on } (-\infty, t]\}.$$

We shall call the *Aubry set of  $u$* :  $\tilde{\mathcal{I}}(u) := \bigcap_{t \geq 0} \tilde{\mathcal{I}}_t(u)$ , that can be also defined as

$$\tilde{\mathcal{I}}(u) := \{(x, v) \in TM : \gamma_{(x, v)}(s) := \pi \Phi_s^L((x, v)) \text{ is } (u, L_\eta, \alpha(c)) - \text{calibr. on } \mathbb{R}\}.$$

These sets  $\tilde{\mathcal{J}}(u)$  are non-empty, compact and invariant. Moreover, here are some properties of these sets (compare with theorem 6.18).

**Proposition 6.20.** *Let  $u \prec L_\eta + \alpha(c)$ .*

- (1)  $\tilde{\mathcal{J}}_t(u)$  is compact;
- (2)  $\tilde{\mathcal{J}}_{t'}(u) \subseteq \tilde{\mathcal{J}}_t(u) \subseteq \tilde{\mathcal{J}}_0(u)$  for all  $t' \geq t \geq 0$ ;
- (3)  $\mathcal{L}(\tilde{\mathcal{J}}_0(u))$  is contained in the energy level  $\mathcal{E}_c^*$  corresponding to  $\alpha(c)$ ;
- (4)  $\mathcal{L}(\tilde{\mathcal{J}}_t(u)) \subseteq \text{Graph}(\eta + du)$  for all  $t > 0$ ;
- (5)  $\mathcal{L}(\tilde{\mathcal{J}}_0(u)) \subseteq \overline{\text{Graph}(\eta + du)}$  (observe that for weak solutions these two sets coincide);
- (6)  $\Phi_{-t}^L(\tilde{\mathcal{J}}_0(u)) = \tilde{\mathcal{J}}_t(u)$  for all  $t > 0$ ;
- (7)  $\bigcup_{t>0} \tilde{\mathcal{J}}_t(u) = \tilde{\mathcal{J}}_0(u)$ ;
- (8)  $\tilde{\mathcal{J}}(u) = \bigcap_{t \geq 0} \Phi_{-t}^L(\tilde{\mathcal{J}}_0(u))$ ;
- (9)  $\tilde{\mathcal{J}}_0(u) \subseteq W^u(\tilde{\mathcal{J}}(u))$ .
- (10)  $\pi : \tilde{\mathcal{J}}(u) \rightarrow \pi(\tilde{\mathcal{J}}(u))$  is a bi-Lipschitz homeomorphism [Graph Theorem].  
The same is true for  $\tilde{\mathcal{J}}_t(u)$  for each  $t > 0$ .

**Theorem 6.21 (Fathi).** *The Aubry and Mañé sets defined in (23) and (22) can be equivalently defined in the following ways:*

$$\begin{aligned} \tilde{\mathcal{A}}_c &= \bigcap_{u \prec L_\eta + \alpha(c)} \tilde{\mathcal{J}}(u) = \bigcap_{u \prec L_\eta + \alpha(c)} \mathcal{L}^{-1}(\text{Graph}(\eta + du)) \\ \tilde{\mathcal{N}}_c &= \bigcup_{u \prec L_\eta + \alpha(c)} \tilde{\mathcal{J}}(u). \end{aligned}$$

Moreover, there exists  $u_\infty \prec L_\eta + \alpha(c)$  such that  $\tilde{\mathcal{A}}_c = \tilde{\mathcal{J}}(u_\infty) \cap \mathcal{L}^{-1}(\mathcal{E}_c^*)$ .

The proof of this theorem follows from the results in [22, Chapter 9]. For the last statement is sufficient to observe that the set of critically dominated functions is a separable subset of  $C(M)$ . Let  $\{u_n\}$  be a countable dense family of such functions and define  $u_\infty$  as a convex combination of their normalization (with respect to a fixed point  $x_0 \in M$ ), e.g.,  $u_\infty(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} (u_n(x) - u_n(x_0))$ .

**Remark 6.22.** Using this characterization, the graph property of the Aubry set (Theorem 5.30) follows easily from property (10) in Proposition 6.20. Moreover, the non-emptiness of  $\tilde{\mathcal{N}}_c$  is a result of the non-emptiness of  $\tilde{\mathcal{J}}(u)$ . As far as the non-emptiness of  $\tilde{\mathcal{A}}_c$  is concerned, one can deduce it from this characterization and proposition 6.33.

From theorem 6.21 one can also deduce another interesting property of critically dominated functions: their differentiability on the projected Aubry set (recall that a-priori these functions are only Lipschitz, so they are differentiable almost everywhere).

**Proposition 6.23** (see [22, Theorem 4.3.8 (i)]). *Let  $u \prec L_\eta + \alpha(c)$ . For each  $x \in \mathcal{A}_c$ ,  $u$  is differentiable at  $x$  and  $d_x u$  does not depend on  $u$ ; namely,  $d_x u = \frac{\partial L}{\partial v}(x, \pi_{\tilde{\mathcal{A}}_c}^{-1}(x))$ .*

In addition to the Aubry set and Mañé set, one can also recover the definition of Mañé potential  $\phi_{\eta, \alpha(c)}$  (see Section 5) and Peierls barrier  $h_\eta$  (see (25)) in terms of these solutions and subsolutions.

Let us start by observing that, from definition 6.3, if  $u \prec L_\eta + \alpha(c)$  then for each  $x, y \in M$  and  $t > 0$  we have that  $u(y) - u(x) \leq h_{\eta,t}(x, y) + \alpha(c)t$ , where  $h_{\eta,t}$  is defined as

$$h_{\eta,t}(x, y) = \min \int_0^t L_\eta(\gamma(s), \dot{\gamma}(s)) ds,$$

where the minimum is taken over all piecewise  $C^1$  paths  $\gamma : [0, t] \rightarrow M$ , such that  $\gamma(0) = x$  and  $\gamma(t) = y$ . This minimum is achieved because of Tonelli theorem (Theorem 5.1).

This implies that for each  $u \prec L_\eta + \alpha(c)$  and for each  $x, y \in M$ ,  $u(y) - u(x) \leq h_\eta(x, y)$ , i.e.,

$$\phi_{\eta, \alpha(c)}(x, y) \geq \sup_{u \prec L_\eta + \alpha(c)} [u(y) - u(x)] \quad \forall x, y \in M.$$

One can actually show that they are equal.

**Proposition 6.24** (see [22, Corollary 9.1.3]). *For each  $x, y \in M$ , we have the equality*

$$\phi_{\eta, \alpha(c)}(x, y) = \sup_{u \prec L_\eta + \alpha(c)} [u(y) - u(x)].$$

**Remark 6.25.** The quantity on the right-hand side is also called “viscosity semi-distance” (see [22, Section 8.4]).

As far as Peierls barrier is concerned, let us observe that similarly to what happens for Mañé potential, also in this case we have that for each  $u \prec L_\eta + \alpha(c)$  and for each  $x, y \in M$ ,  $u(y) - u(x) \leq h_\eta(x, y)$ , i.e.,

$$h_\eta(x, y) \geq \sup_{u \prec L_\eta + \alpha(c)} [u(y) - u(x)] \quad \forall x, y \in M.$$

Moreover, if  $u_- \in \mathcal{S}_\eta^-$  and  $u_+ \in \mathcal{S}_\eta^+$  are conjugate solutions, the same result holds:  $u_-(y) - u_+(x) \leq h_\eta(x, y)$  and consequently

$$h_\eta(x, y) \geq \sup_{(u_-, u_+)} [u_-(y) - u_+(x)] \quad \forall x, y \in M,$$

where  $(u_-, u_+)$  denotes conjugate weak KAM solutions. In addition to this, it is possible to show that the above inequality is actually an equality. In fact:

**Proposition 6.26** (see [18, Théorème 7]). *For  $x \in M$  let us define the function  $h_\eta^x : M \rightarrow \mathbb{R}$  (resp.  $h_{\eta,x} : M \rightarrow \mathbb{R}$ ) by  $h_\eta^x(y) = h_\eta(x, y)$  (resp.  $h_{\eta,x}(y) = h_\eta(y, x)$ ). For each  $x \in M$ , the function  $h_\eta^x$  (resp.  $-h_{\eta,x}$ ) is a weak KAM solution of negative (resp. positive) type. Moreover, its conjugate function  $u_+^x \in \mathcal{S}_\eta^+$  (resp.  $u_-^x \in \mathcal{S}_\eta^-$ ) vanishes at  $x$ .*

Therefore:

**Corollary 6.27.** *For each  $x, y \in M$ , we have the equality*

$$h_\eta(x, y) = \sup_{(u_-, u_+)} [u_-(y) - u_+(x)],$$

where the supremum is taken over pairs of conjugate solutions. Moreover, for any given  $x, y \in M$  this supremum is actually attained.

Observe that, since for any  $u \prec L_\eta + \alpha(c)$  there exists a weak KAM solution of negative type  $u_-$  and a weak KAM solution of positive type  $u_+$ , such that  $u_- = u = u_+$  on the projected Aubry set  $\mathcal{A}_c$  (see Theorem 6.16), one can get the following representations for Peierls barrier  $h_\eta$  on the projected Aubry set:

$$(29) \quad h_\eta(x, y) = \sup_{u \prec L_\eta + \alpha(c)} [u(y) - u(x)]$$

for all  $x, y \in \mathcal{A}_c$ . This supremum is actually attained for any fixed  $x \in \mathcal{A}_c$ .

**Remark 6.28.** In particular, (29) shows that Peierls barrier and Mañé potential coincide on the projected Aubry set.

## ADDENDUM

### 6.A - REGULARITY OF CRITICAL SUBSOLUTIONS

In this addendum we want to say more about the regularity of critically dominated functions or  $\eta$ -critical subsolutions.

We have remarked above in this section, that for  $k < \alpha(c)$  there do not exist functions dominated by  $L_\eta + k$ , while for  $k \geq \alpha(c)$  they do exist. Moreover, if  $k > \alpha(c)$  these functions can be chosen to be  $C^\infty$  (see also characterization of  $\alpha(c)$  in remark 6.8). The critical case has totally different features. As a counterpart of their relation with the dynamics of the system, critical dominated functions have very rigid structural properties, that become an obstacle when someone tries to make them smoother. For instance, as we have recalled in proposition 6.23, if  $u \prec L_\eta + \alpha(c)$  then its differential  $d_x u$  exists on  $\mathcal{A}_c$  and it is prescribed over there. This means that although it is quite easy to make these functions smoother (e.g.,  $C^\infty$ ) out of the projected Aubry set, it is impossible to modify them on this set.

Nevertheless, Albert Fathi and Antonio Siconolfi [26] managed to prove that  $C^1$   $\eta$ -critical subsolutions do exist and are dense, in the following sense:

**Theorem 6.29 (Fathi, Siconolfi).** *Let  $u \prec L_\eta + \alpha(c)$ . For each  $\varepsilon > 0$ , there exists a  $C^1$  function  $\tilde{u} : M \rightarrow \mathbb{R}$  such that:*

- i)  $\tilde{u} \prec L_\eta + \alpha(c)$ ;
- ii)  $\tilde{u}(x) = u(x)$  on  $\mathcal{A}_c$ ;
- iii)  $|\tilde{u}(x) - u(x)| < \varepsilon$  on  $M \setminus \mathcal{A}_c$ .

Moreover, one can choose  $\tilde{u}$  so that it is a strict  $\eta$ -critical subsolution, i.e., we have  $H_\eta(x, d_x \tilde{u}) < a(c)$  on  $M \setminus \mathcal{A}_c$ .

This result has been extended by Patrick Bernard [6], who showed that every  $\eta$ -critical subsolution coincides, on the Aubry set, with a  $C^{1,1}$   $\eta$ -critical subsolution.

**Theorem 6.30 (Bernard, [6]).** *Let  $H$  be a Tonelli Hamiltonian. If the Hamilton-Jacobi equation has a subsolution, then it has a  $C^{1,1}$  subsolution. Moreover, the set of  $C^{1,1}$  subsolutions is dense for the uniform topology in the set of subsolutions.*

**Remark 6.31.** In general  $C^{1,1}$  is the best regularity that one can expect: it is easy in fact to construct examples in which  $C^2$   $\eta$ -critical subsolutions do not exist. For example, consider the case in which the Aubry set projects over all the manifold  $M$  and it is not a  $C^1$  graph (e.g., on  $M = \mathbb{T}$  take  $L(x, v) = \frac{1}{2}\|v\|^2 + \sin^2(\pi x)$  and  $\eta = \frac{2}{\pi}dx$ ). In this case there is only one critical subsolution (up to constants), that is an actual solution: its differential is Lipschitz but not  $C^1$ .

It is therefore clear that the structure of the Aubry set plays a crucial role. Patrick Bernard [7] proved that if the Aubry set is a union of finitely many hyperbolic periodic orbits or hyperbolic fixed points, then smoother subsolutions can be constructed. In particular, if the Hamiltonian is  $C^k$ , then these subsolutions will be  $C^k$  too. The proof of this result is heavily based on the hyperbolic structure of the Aubry set and the result is deduced from the regularity of its local stable and unstable manifolds.

See also [21] for a survey on this problem and some results concerning Denjoy-type obstructions for the existence of regular critical subsolutions on  $\mathbb{T}^2$ .

Using the density of  $C^1$  critically dominated functions, one can improve some of the results in (28), Theorem 6.21 and (29). Let us denote by  $\mathcal{S}_\eta^1$  the set of  $C^1$   $\eta$ -critical subsolutions and  $\mathcal{S}_\eta^{1,1}$  the set of  $C^{1,1}$   $\eta$ -critical subsolutions. Then:

$$\begin{aligned} \tilde{\mathcal{A}}_c &:= \bigcap_{u \in \mathcal{S}_\eta^1} \tilde{\mathcal{J}}(u) = \bigcap_{u \in \mathcal{S}_\eta^{1,1}} \tilde{\mathcal{J}}(u) = \\ (30) \quad &= \bigcap_{u \in \mathcal{S}_\eta^1} \mathcal{L}^{-1}(\text{Graph}(\eta + du)) = \bigcap_{u \in \mathcal{S}_\eta^{1,1}} \mathcal{L}^{-1}(\text{Graph}(\eta + du)). \end{aligned}$$

In particular, there exists a  $C^{1,1}$   $\eta$ -critical subsolution  $\tilde{u}$  such that:

$$\begin{aligned} \tilde{\mathcal{A}}_c &= \mathcal{L}^{-1}(\text{Graph}(\eta + d\tilde{u}) \cap \mathcal{E}_c^*) = \\ (31) \quad &= \tilde{\mathcal{J}}(\tilde{u}) \cap \mathcal{L}^{-1}(\mathcal{E}_c^*). \end{aligned}$$

Moreover, for any  $x, y \in \mathcal{A}_c$ :

$$(32) \quad h_\eta(x, y) = \sup_{u \in \mathcal{S}_\eta^1} \{u(y) - u(x)\} = \sup_{u \in \mathcal{S}_\eta^{1,1}} \{u(y) - u(x)\}$$

where all the above suprema are maxima for any fixed  $x, y \in \mathcal{A}_c$ .

## 6.B - NON-WANDERING SET OF THE MAÑÉ SET

In this addendum we want now to use this approach and the above-mentioned results to show that the non-wandering set of the Euler-Lagrange flow restricted to the Mañé set is contained in the Aubry set (we mentioned this in Proposition 5.28). Let us first recall the definition of *non-wandering point* for a flow  $\Phi_t : X \rightarrow X$ .

**Definition 6.32.** *A point  $x \in X$  is called non-wandering if for each neighborhood  $\mathcal{U}$  and each positive integer  $n$ , there exists  $t > n$  such that  $f^t(\mathcal{U}) \cap \mathcal{U} \neq \emptyset$ .*

We shall denote the set of *non-wandering points* for  $\Phi_t$  by  $\Omega(\Phi_t)$ .

**Proposition 6.33.** *If  $M$  is a compact manifold and  $L$  a Tonelli Lagrangian on  $TM$ , then  $\Omega(\Phi_t^L|_{\tilde{\mathcal{N}}_c}) \subseteq \tilde{\mathcal{A}}_c$  for each  $c \in H^1(M; \mathbb{R})$ .*

### Remarks.

- 1) Proposition 6.33 also shows that the Aubry set is non-empty. In fact, any continuous flow on a compact space possesses non-wandering points.
- 2) Since every point in the support of an invariant measure  $\mu$  is non-wandering, then this also shows that  $\tilde{\mathcal{M}}_c \subseteq \tilde{\mathcal{A}}_c$ .

**Proof.** Let  $(x, v) \in \Omega(\Phi_t^L|_{\tilde{\mathcal{N}}_c})$ . By the definition of non-wandering point, there exist a sequence  $(x_k, v_k) \in \tilde{\mathcal{N}}_c$  and  $t_k \rightarrow +\infty$ , such that  $(x_k, v_k) \rightarrow (x, v)$  and  $\Phi_{t_k}^L(x_k, v_k) \rightarrow (x, v)$  as  $k \rightarrow +\infty$ . From (30), for each  $(x_k, v_k)$  there exists a  $\eta$  critical subsolution  $u_k$ , such that the curve  $\gamma_k(t) = \pi(\Phi_t^L(x_k, v_k))$  is  $(u_k, L_\eta, \alpha(c))$ -calibrated. Moreover, up to extracting a subsequence, we can assume that, on any compact interval,  $\gamma_k$  converge in the  $C^1$ -topology to  $\gamma(t) = \pi(\Phi_t^L(x, v))$ .

Pick now any critical subsolution  $u$ . If we show that  $\gamma$  is  $(u, L_\eta, \alpha(c))$ -calibrated, using (30) we can conclude that  $(x, v) \in \tilde{\mathcal{A}}_c$ . First of all, observe that, by the continuity of  $u$ ,

$$u(\gamma_k(t_k)) - u(x_k) \xrightarrow{k \rightarrow \infty} 0.$$

Using that  $\eta$ -critical subsolutions are equi-Lipschitz (as remarked after definition 6.3), we can also conclude that

$$u_k(\gamma_k(t_k)) - u_k(x_k) \xrightarrow{k \rightarrow \infty} 0$$

and, therefore,

$$(33) \quad \int_0^{t_k} L_\eta(\gamma_k(s), \dot{\gamma}_k(s)) + \alpha(c) ds = u_k(\gamma_k(t_k)) - u_k(x_k) \xrightarrow{k \rightarrow \infty} 0.$$

Let  $0 \leq a \leq b$  and choose  $t_k \geq b$ . Observe now that

$$\begin{aligned} u(\gamma_k(b)) - u(\gamma_k(a)) &= u(\gamma_k(t_k)) - u(x_k) - [u(\gamma_k(t_k)) - u(\gamma_k(b))] - \\ &\quad - [u(\gamma_k(a)) - u(x_k)] \geq \\ &\geq u(\gamma_k(t_k)) - u(x_k) - \int_b^{t_k} L_\eta(\gamma_k(s), \dot{\gamma}_k(s)) + \alpha(c) ds - \\ &\quad - \int_0^a L_\eta(\gamma_k(s), \dot{\gamma}_k(s)) + \alpha(c) ds = \\ &= u(\gamma_k(t_k)) - u(x_k) + \int_a^b L_\eta(\gamma_k(s), \dot{\gamma}_k(s)) + \alpha(c) ds - \\ &\quad - \int_0^{t_k} L_\eta(\gamma_k(s), \dot{\gamma}_k(s)) + \alpha(c) ds; \end{aligned}$$

taking the limit as  $k \rightarrow \infty$  on both sides, one can conclude:

$$u(\gamma(b)) - u(\gamma(a)) \geq \int_a^b L_\eta(\gamma(s), \dot{\gamma}(s)) + \alpha(c) ds$$

and therefore, from the fact that  $u \prec L_\eta + \alpha(c)$ , it follows the equality. This shows that  $\gamma$  is  $(u, L_\eta, \alpha(c))$ -calibrated on  $[0, \infty)$ . To show that it is indeed calibrated on all  $\mathbb{R}$ , one can make a symmetric argument, letting  $(y_k, w_k) = \Phi_{t_k}^L(x_k, v_k)$  play the role of  $(x_k, v_k)$  in the previous argument. In fact, one has  $(y_k, w_k) \rightarrow (x, v)$  and  $\Phi_{-t_k}^L(y_k, w_k) \rightarrow (x, v)$  as  $k \rightarrow +\infty$  and the very same argument works.  $\square$

## REFERENCES

- [1] Vladimir I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18(5 (113)):13–40, 1963.
- [2] Vladimir I. Arnol'd. Instability of dynamical systems with many degrees of freedom. *Dokl. Akad. Nauk SSSR*, 156:9–12, 1964.
- [3] Serge Aubry and P. Y. Le Daeron. The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D*, 8(3):381–422, 1983.
- [4] J. M. Ball and V. J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. *Arch. Rational Mech. Anal.*, 90(4):325–388, 1985.
- [5] Patrick Bernard. Connecting orbits of time dependent Lagrangian systems. *Ann. Inst. Fourier (Grenoble)*, 52(5):1533–1568, 2002.
- [6] Patrick Bernard. Existence of  $C^{1,1}$  critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds. *Ann. Sci. École Norm. Sup. (4)*, 40(3):445–452, 2007.
- [7] Patrick Bernard. Smooth critical sub-solutions of the Hamilton-Jacobi equation. *Math. Res. Lett.*, 14(3):503–511, 2007.
- [8] Patrick Bernard. Symplectic aspects of Mather theory. *Duke Math. J.*, 136(3):401–420, 2007.
- [9] Patrick Bernard. The dynamics of pseudographs in convex Hamiltonian systems. *J. Amer. Math. Soc.*, 21(3):615–669, 2008.
- [10] Patrick Bernard. Young measures, superposition and transport. *Indiana Univ. Math. J.*, 57(1):247–275, 2008.
- [11] Patrick Bernard and Gonzalo Contreras. A generic property of families of Lagrangian systems. *Ann. of Math. (2)*, 167(3):1099–1108, 2008.

- [12] Patrick Bernard and Joana Santos. A geometric definition of the Aubry-Mather set. Preprint, 2010.
- [13] Mario J. Dias Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity*, 8(6):1077–1085, 1995.
- [14] Gonzalo Contreras, Jorge Delgado, and Renato Iturriaga. Lagrangian flows: the dynamics of globally minimizing orbits. II. *Bol. Soc. Brasil. Mat. (N.S.)*, 28(2):155–196, 1997.
- [15] Gonzalo Contreras, Renato Iturriaga, Gabriel P. Paternain, and Miguel Paternain. Lagrangian graphs, minimizing measures and Mañé's critical values. *Geom. Funct. Anal.*, 8(5):788–809, 1998.
- [16] Gonzalo Contreras, Leonardo Macarini, and Gabriel P. Paternain. Periodic orbits for exact magnetic flows on surfaces. *Int. Math. Res. Not.*, (8):361–387, 2004.
- [17] Gonzalo Contreras and Gabriel P. Paternain. Connecting orbits between static classes for generic Lagrangian systems. *Topology*, 41(4):645–666, 2002.
- [18] Albert Fathi. Solutions KAM faibles conjuguées et barrières de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(6):649–652, 1997.
- [19] Albert Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1043–1046, 1997.
- [20] Albert Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(3):267–270, 1998.
- [21] Albert Fathi. Denjoy-Schwartz and Hamilton-Jacobi. *J. Math. Kyoto Univ.* (To appear), 2009.
- [22] Albert Fathi. Weak KAM theorem and Lagrangian dynamics. 10th Preliminary version, 2009.
- [23] Albert Fathi, Alessio Figalli, and Ludovic Rifford. On the Hausdorff dimension of the Mather quotient. *Comm. Pure Appl. Math.*, 62(4):445–500, 2009.
- [24] Albert Fathi, Alessandro Giuliani, and Alfonso Sorrentino. Uniqueness of invariant Lagrangian graphs in a homology class or a cohomology class. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, VIII(4):659–680, 2009.
- [25] Albert Fathi and John N. Mather. Failure of convergence of the Lax-Oleinik semi-group in the time-periodic case. *Bull. Soc. Math. France*, 128(3):473–483, 2000.
- [26] Albert Fathi and Antonio Siconolfi. Existence of  $C^1$  critical subsolutions of the Hamilton-Jacobi equation. *Invent. Math.*, 155(2):363–388, 2004.
- [27] Giovanni Forni and John N. Mather. Action minimizing orbits in Hamiltonian systems. In *Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991)*, volume 1589 of *Lecture Notes in Math.*, pages 92–186. Springer, Berlin, 1994.
- [28] Gustav A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. of Math. (2)*, 33(4):719–739, 1932.
- [29] Michael R. Herman. Inégalités “a priori” pour des tores lagrangiens invariants par des difféomorphismes symplectiques. *Inst. Hautes Études Sci. Publ. Math.*, (70):47–101 (1990), 1989.
- [30] Andrey N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.
- [31] Nicolas Kryloff and Nicolas Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Ann. of Math. (2)*, 38(1):65–113, 1937.
- [32] Pierre-Louis Lions, George C. Papanicolau, and Srinivasa R. S. Varadhan. Homogenization of Hamilton-Jacobi equation. Unpublished preprint, 1987.
- [33] Ricardo Mañé. On the minimizing measures of Lagrangian dynamical systems. *Nonlinearity*, 5(3):623–638, 1992.
- [34] Ricardo Mañé. Generic properties and problems of minimizing measures of Lagrangian systems. *Nonlinearity*, 9(2):273–310, 1996.
- [35] Ricardo Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bol. Soc. Brasil. Mat. (N.S.)*, 28(2):141–153, 1997.
- [36] Daniel Massart. Normes stables des surfaces. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(2):221–224, 1997.
- [37] Daniel Massart. On Aubry sets and Mather's action functional. *Israel J. Math.*, 134:157–171, 2003.
- [38] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21(4):457–467, 1982.
- [39] John N. Mather. More Denjoy minimal sets for area preserving diffeomorphisms. *Comment. Math. Helv.*, 60(4):508–557, 1985.
- [40] John N. Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, (63):153–204, 1986.

- [41] John N. Mather. Modulus of continuity for Peierls's barrier. In *Periodic solutions of Hamiltonian systems and related topics (Il Ciocco, 1986)*, volume 209 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 177–202. Reidel, Dordrecht, 1987.
- [42] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.
- [43] John N. Mather. Variational construction of orbits of twist diffeomorphisms. *J. Amer. Math. Soc.*, 4(2):207–263, 1991.
- [44] John N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)*, 43(5):1349–1386, 1993.
- [45] John N. Mather. Arnol'd diffusion. I. Announcement of results. *Sovrem. Mat. Fundam. Napravl.*, 2:116–130 (electronic), 2003.
- [46] John N. Mather. Total disconnectedness of the quotient Aubry set in low dimensions. *Comm. Pure Appl. Math.*, 56(8):1178–1183, 2003. Dedicated to the memory of Jürgen K. Moser.
- [47] John N. Mather. Examples of Aubry sets. *Ergodic Theory Dynam. Systems*, 24(5):1667–1723, 2004.
- [48] John N. Mather. Arnol'd diffusion. II. Unpublished Manuscript, 2007.
- [49] Jürgen Moser. Monotone twist mappings and the calculus of variations. *Ergodic Theory Dynam. Systems*, 6(3):401–413, 1986.
- [50] Gabriel P. Paternain, Leonid Polterovich, and Karl Friedrich Siburg. Boundary rigidity for Lagrangian submanifolds, non-removable intersections, and Aubry-Mather theory. *Mosc. Math. J.*, 3(2):593–619, 745, 2003.
- [51] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [52] Sol Schwartzman. Asymptotic cycles. *Ann. of Math. (2)*, 66:270–284, 1957.
- [53] Alfonso Sorrentino. On the total disconnectedness of the quotient Aubry set. *Ergodic Theory Dynam. Systems*, 28(1):267–290, 2008.
- [54] Alfonso Sorrentino. On the integrability of Tonelli Hamiltonians. *Preprint*, 2009. To appear on *Trans. Amer. Math. Soc.*

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE, WILBERFORCE ROAD, CAMBRIDGE CB3 0WB, UNITED KINGDOM  
*E-mail address:* as998@dpms.cam.ac.uk